ANALYSES OF SPEECH PROCESSING STRATEGIES FOR COCHLEAR IMPLANTS AND THE EFFECTS OF ELECTRODE INTERACTION

APPROVED BY SUPERVISORY COMMITTEE:

Dr. Peter F. Assmann, Chair

Dr. Philip C. Loizou

Dr. Linda Thibodeau

Dr. Emily Tobey

ACKNOWLEDGEMENTS

No study or investigation of this magnitude is ever completed by the efforts of a single person. I am forever grateful to the patients for their contribution to science and my dissertation. I would like to thank my supervising professor, Dr. Peter Assmann, for his invaluable guidance, teaching and encouragement during the writing of this dissertation and throughout the entire duration of my time spent as a graduate student here at UT – Dallas. I would also like to thank Dr. Philip Loizou and Dr. Bob Shannon who inspired this project, and for their valuable insight and guidance on my dissertation.

ABSTRACT

Multichannel cochlear implants electrically stimulate the auditory nerve to restore partial hearing to the profoundly deaf patient. The multichannel implant was designed to selectively stimulate discrete populations of spiral ganglion cells along the length of the cochlea. However, selective stimulation is not often, or at least imperfectly, achieved even with the most modern cochlear implant designs and speech processing strategies. When multiple electrodes are stimulated simultaneously, electrical fields generated around each electrode can interact with the electrical fields of neighboring electrodes, thereby reducing selectivity. Several studies have suggested that electrical-field interactions can disrupt the acoustic properties of the signal and severely degrade speech intelligibility, however this relationship has not been directly tested.

Electrical-field interactions can be reduced by decreasing the current levels delivered to each electrode through improved electrode positioning and design, or by using speech processing strategies that maximize the separation between simultaneously stimulated electrodes or stimulate the electrodes sequentially. The proximity of the cochlear implant electrode array to the modiolus has been shown to reduce the amount of current required to reach threshold (Rebscher et al., 1994). When less current is required, current spread and electrical field overlap is reduced. Recently, cochlear implant manufacturers have taken interest in designing “positioners” which place the electrode array in close proximity to the spiral ganglion cells and new electrode arrays which attempt to direct their current toward spiral ganglion cell bodies.

The following experiments examine electrical-field interactions and speech recognition performance for three electrode designs: patients implanted with the Enhanced Bipolar Clarion electrode array without a “positioner”, patients implanted with the Clarion Electrode Positioning System™ (EPS) and the Enhanced Bipolar electrode array, and patients with the EPS and the Clarion Hi-Focus™ electrode array. A simultaneous masking task was used to measure electrical-field interactions as a function of electrode separation for monopolar and bipolar configurations. The relationship between electrical-field interaction and speech recognition was also examined for several speech strategies varying in the number of electrodes stimulated simultaneously. Subjects identified consonants, vowels, and sentences with each of the following speech strategies, listed in order from sequential stimulation to fully simultaneous stimulation: Continuous Interleaved Sampler (CIS), Paired Pulsatile Sampler (PPS), Quadruple Pulsatile Sampler (QPS), Hybrid Analog Pulsatile (HAPs), and Simultaneous Analog Stimulation (SAS). Based on previous research, susceptibility to electrical-field interactions is expected to vary as a function of electrode design, the speech processing strategy used in the device, and factors specific to each patient. The contribution from each of these variables was investigated.

The results showed a moderate to strong negative correlation between electrical-field interaction and speech recognition performance, which indicates that patients with lower levels of electrical-field interaction have higher speech recognition scores than patients with high levels of electrical-field interaction. In addition, patients with strong susceptibility to electrical-field interactions produced higher speech recognition scores

for sequential than simultaneous speech strategies. An information analysis revealed that vowel recognition and consonant place-of-articulation were most affected by electrical-field interactions, demonstrating that electrode interactions severely disrupt spectral cues. The pattern of results also suggests that, with acute listening trials, patients achieve the highest speech recognition scores with the speech processing strategy most similar to their own. Future studies are needed to determine if patients with minimal levels of electrical-field interaction can benefit from the partially-simultaneous QPS or HAPs strategies with more listening exposure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

General Introduction

The cochlear implant is the only available medical device that can restore partial hearing to patients with profound deafness. The implant bypasses the damaged region of the inner ear and stimulates the auditory nerve directly with an electrical representation of the acoustic signal. Currently there are over 30,000 cochlear implant users worldwide, with some being able to carry on a conversation via telephone. However, even with the same technology and hearing history, cochlear implant users show an extremely wide range of performance variability. Accounting for these performance differences is a central topic in cochlear implant research. An increased understanding of the basic perceptual mechanisms employed by CI users could lead to better speech processing strategies and optimized clinical fitting protocols. Knowledge of how listeners extract auditory meaning from a distorted or impoverished signal might also reveal some of the perceptual and cognitive mechanisms used by listeners with normal hearing.

Several researchers have suggested that electrical-field interaction, i.e. the summation of current that can occur with simultaneous electrode stimulation, may account for at least some of the performance variability (Hanekom and Shannon, Shannon, 1985; White, Merzenich, & Gardi, 1984; Wilson et al, 1991). However, to date, speech recognition performance and electrical-field interactions have never been evaluated and related in the same study.

One means of reducing electrical-field interactions is by developing new speech processing strategies and modifying the parameters of existing strategies. Speech processing strategies are the programs that determine how the electrodes should be stimulated, i.e. the rate or order of stimulation. The range of speech processing strategies available to cochlear implant patients has grown considerably in recent years. Several manufacturers have developed multiple strategies for their device. Since no one strategy is optimal for all patients, having the option to choose among the various strategies and their corresponding parameters (i.e. pulse rate, stimulation order, coupling mode, number of simultaneous channels, etc.) may raise the average level of speech recognition performance.

However, electrical-field interactions, arising from simultaneous electrode stimulation, can severely restrict the range of strategies that provide intelligible speech. It has been reported that approximately 1/3 of Clarion patients prefer the Simultaneous Analog Sampler (SAS) strategy (Armstrong-Bednall et al., 1999; Osberger & Fisher, 1999; Stollwerck et al., 1999). Ideally, all patients should benefit from speech processing strategies like SAS, which have the capability of delivering the most information because of their simultaneous nature. As a result, cochlear implant researchers and implant manufacturers continue to explore techniques on how to best modify speech processing strategies and the implant hardware to reduce the chance for electrical-field interactions.

With the recent development of a signal generator called the Clarion Research Interface (CRI), researchers can now implement and modify the parameters of a wide range of speech processing strategies. Several existing and new speech strategies were developed with the CRI and tested in this study. Each of the speech strategies represented a range of stimulation from purely simultaneous to sequential electrode stimulation. It was hypothesized that speech recognition performance along this continuum of speech strategies would show a strong relationship with the degree of electrical-field interaction experienced by that patient. Patients with more electrical-field interaction were predicted to show higher speech recognition performance with sequential, or non-simultaneous, speech strategies (e.g. CIS), and patients with very little to no electrical-field interaction were predicted to prefer and achieve higher speech recognition scores with simultaneous speech strategies (e.g. SAS). Knowing the point at which electrical interactions are minimized could be invaluable in the design of future speech processing strategies and electrode arrays.

Chapter 2 presents data from a pilot study where electrical-field interactions were measured between adjacent electrodes at several points along the cochlea. Electrical-field interactions were compared for three Clarion electrode configurations to evaluate their effectiveness at reducing electrical-field interactions. The relationship between "local" electrical-field interactions, as measured between adjacent electrodes, and three speech recognition tasks (vowels, consonants, and sentences) is discussed.

In Chapter 3, data is presented from experiments measuring electrical-field interactions as a function of electrode separation. The electrode separation necessary for electrical-field interactions to drop to negligible levels was investigated for the three Clarion electrode configurations evaluated in the pilot study. This information can shed some light on the patient's preference for a particular speech processing strategy.

In Chapter 4, speech recognition performance was evaluated during acute listening sessions for five speech processing strategies. Each of the strategies varied in the number of simultaneous electrodes. It was hypothesized that only patients with low levels of electrical-field interactions (as determined from the experiments in Chapter 3) would have high speech recognition scores with as the number of simultaneous channels used by the speech strategy increased.

In Chapter 5, the relationship between speech recognition performance and electrical-field interactions is examined. The contribution from other factors, such as duration of deafness, to speech recognition performance is also investigated.

CHAPTER 2

Literature Review

One of the longstanding issues in hearing science is the nature and importance of cochlear frequency selectivity for speech understanding. Cochlear implant research can be used to explore these basic hearing science questions since successive improvements in cochlear implant design are believed to be the result of more detailed modeling of normal cochlear function.

The experiments in this thesis examine speech recognition in patients who have developed oral language prior to their hearing loss. For this reason, the following literature review will limit the discussion to cochlear implantation in adult, postlingually-deafened listeners.

2.1 The Auditory Coding of Sound

The information-bearing components of the speech signal include both amplitude and frequency variation over time. Frequency discrimination can be achieved by the peripheral auditory system with either a rate-place or phase-locking representation, or a combination of both (Moller, 1999).

According to the rate-place representation, sound is coded by the relative distribution of neural firing patterns as a function of signal frequency: high frequency pure tones elicit a higher rate of firing from auditory nerve fibers at the base of the cochlea and nerve fibers at the apex respond best to low frequency tones. The auditory nerve fibers within the cochlea are therefore described as having a tonotopic organization. Several studies have shown that the tonotopical representation is maintained, and possibly even enhanced, in the auditory cortex (Aitkin, 1990).

Neural phase-locking is another mechanism with which speech information can be coded by the auditory system. An incoming sound wave sets the basilar membrane into motion. Each auditory nerve fiber has a greater *probability* of responding to a particular phase of vibration generated by the stimulating waveform. The composite response from an ensemble of nerve fibers can relay a fairly accurate representation of the waveform.

2.2 Single Channel Implants

The first generation of cochlear implants, known as single-channel implants, avoided cochlear frequency analysis, or place coding, altogether. Single-channel implants emphasized the temporal-amplitude variations of the speech waveform by either increasing or decreasing the current level to represent the sound intensity fluctuations of speech over time (Figure 2.2, top trace). The top trace in the figure below shows the broadband amplitude envelope superimposed on the waveform of the sentence "The watchdog gave a warning growl." and roughly represents the speech information coded by a single-channel implant.

Amplitude Envelope of a Sentence

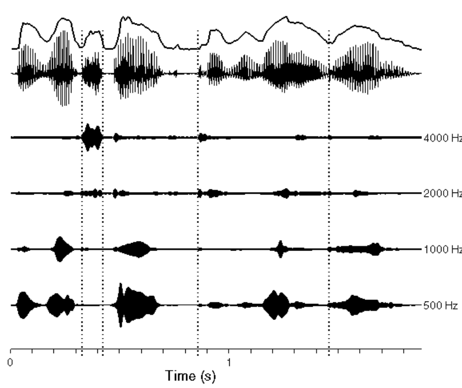


Figure 2.2: The top trace shows the broadband amplitude envelope of the sentence "The watchdog gave a warning growl." and roughly represents the speech information coded by a single-channel implant. The lower four traces are bandpass filtered waveforms of the same sentence with center frequencies of 4000, 2000, 1000, and 500 Hz. These lower traces approximate the bandpass waveforms delivered to the electrodes in a multichannel

> implant (see Section 2.4), where each electrode receives one of the waveforms according to the electrode's place in the cochlea. High frequency waveforms are delivered to electrodes near the base of the cochlea and low frequency waveforms are delivered to apical electrodes.

Although limited, speech perception with a single-channel device was possible with only temporal and amplitude cues, provided the patient received both auditory and lipreading cues (Gantz et al., 1988; Tye-Murray and Tyler, 1989). However, there have been a limited number of studies showing that speech recognition can occur without lipreading (Berliner and Eisenberg, 1987; Berliner et al., 1989a, 1989b). In the Berliner et al. study (1989a) fifty-two percent of the patients had some open-set speech understanding with the single channel device. Their scores for an auditory-only word recognition task were on average 33.3% with a range from 8.3% to 75%. The average level of sentence comprehension was 41.5%. These values are much higher than would be obtained by chance, which is 0% for open-set speech understanding. A simulation of a single-channel implant was used in a study by Van Tasell et al. (1987,1992) where normal hearing listeners were presented with white noise modulated by the envelope of the speech waveform. In a study by Rosen et al., 1989, patients with a single-channel implant were tested and their results compared with the normal hearing listeners from Van Tasell et al., 1987. It was concluded that both groups of listeners exhibited similar performance levels and categorized consonant stimuli in the same way. Specifically, they had little difficulty extracting consonant information with a heavy temporal emphasis (i.e. consonant manner and voicing information), but were unable to accurately extract consonant information with a spectral emphasis (i.e. consonant place information). This is covered in more detail in the following section.

2.3 Acoustic Cues for Speech Recognition in Cochlear Implants

Djourno and Eyries (1957) were the first to attempt direct electrical stimulation to the auditory nerve when an electrode was temporarily implanted on the VIII nerve during surgery for cholesteatoma of a bilaterally deafened patient. The patient reported that the sounds produced by the electrode stimulation resembled a spinning roulette wheel or chirping crickets, depending on the rate of stimulation. Although the patient was unable to understand speech, he did report that the rhythm of the electrically transmitted signal aided him in lipreading.

Psychophysical studies with cochlear implant patients have demonstrated that the rate of stimulation and the corresponding change in the perceived pitch (i.e. “rate pitch”) occurs for rates up to 300-500 Hz - the pitch increases with increasing stimulation rate. The pitch begins to saturate at stimulation frequencies of approximately 500 Hz (Burns and Viemeister, 1976, 1981; Eddington et al, 1978; Pfingst, 1988; Shannon, 1983a). Frequency coding by a single electrode is therefore limited to the frequency of the first formant and the fundamental frequency, or its perceptual correlate, voice pitch. Information about the pitch and first formant when combined with temporal envelope cues can provide information about consonant manner, the presence or absence of consonant voicing, syllable number, and sentence prosody (Dorman, 1993; Erber, 1979; Rosen, 1989; Shannon et al., 1995; Van Tasell et al. 1987, 1992). However, the identification of vowel and consonant place-of-articulation depends on the spectra, or the relative energy across several frequency bands, and the movement of energy across those bands, representing formant transitions. A single electrode cannot encode sufficient spectral detail to adequately define these speech features since there can be no simultaneous representation of different frequencies across the speech frequency range.

2.4 Multichannel Implants

Multichannel implants were introduced in the late 1970s to overcome the limitations of frequency coding found with single-channel implants. Multichannel implants provide both a temporal-amplitude and spectral representation of an auditory event by using a distributed array of electrodes to selectively stimulate discrete populations of spiral ganglion cells at multiple sites in the cochlea (i.e. frequency “place” coding). The incoming signal is divided into several contiguous frequency bands, high frequency signals are delivered to basal electrodes while low-frequency signals are delivered to apical electrodes (see Figure 2.2). The lower traces in the figure are bandpass filtered waveforms with center frequencies of 4000, 2000, 1000, and 500 Hz.

The rationale for using this stimulation method stems from classic physiological studies which have demonstrated that single auditory nerve fibers at the base of the cochlea show a high frequency selectivity, while apical fibers respond best to low frequencies. With more complex signals, such as speech, there are many frequency components continuously changing in intensity and duration. Speech coding is therefore based on the activity of thousands of nerve fibers whose center frequencies (CF) span a broad range of frequency selectivity.

A ground-breaking study by Sachs and Young (1979) showed that the gross spectral envelope of a steady-state vowel can be approximated by the mean firing rate of the tonotopically distributed nerve fibers, with each speech stimulus eliciting a distinct neural excitation pattern. However, they also showed that, at moderate and high sound levels, the spectral representation of the formant peaks became less distinct as the fibers with CFs not matching the formant frequencies increased their discharge rate while the fibers at the formant frequencies reached their saturation point.

On the basis of these findings, it is difficult to envision how spectral information can be reliably encoded at sound levels typical of conversational speech. This problem might be overcome with low spontaneous rate, high threshold fibers, which have a much wider dynamic range (Blackburn and Sachs, 1990; Sachs, Winslow, and Blackburn, 1988; Sachs and Young, 1980; Young and Sachs, 1979) and therefore do not saturate as easily to the phase of the signal. In addition, the auditory coding could be enhanced in the cochlear nucleus and higher auditory centers (Blackburn and Sachs, 1990). Another possibility considered by Young and Sachs (1979) is that the temporal information provided by phase-locking across the tonotopic array might also convey information about the speech spectrum to higher centers.

2.5 Importance of Spectral Selectivity in Cochlear Implants

There are several lines of cochlear implant research that indicate that the central auditory system requires at least a crude spectral pattern for understanding speech. With the introduction of the multichannel implant, patients were beginning to demonstrate that they could understand speech without the aid of lipreading (Chouard and MacLeod, 1976). Patients initially implanted with the single-channel device and then later re-implanted with the multichannel implant clearly showed an improvement in speech perception performance (adults: Rubinstein et al., 1998; children: Miyamoto et al., 1994).

Several aspects of the nature and extent of spectral coding are actively being addressed in the cochlear implant literature. The following questions have been raised: Is it necessary to match the analysis filters of the implant with the tonotopic location of the electrode? How many electrodes are required to reach high levels of speech recognition? To address the first question, it has been reported that the frequency allocation for each electrode in a multichannel implant is important for speech perception. In other words, the electrode should transmit frequencies that match the tonotopic location on the cochlea. Rabinowitz and Eddington (1995) demonstrated that a patient with a multichannel implant exhibited even poorer speech recognition scores with the reversed mapping (i.e. high frequencies delivered to apical electrodes and low frequencies to basal electrodes) than with a single-channel mapping (i.e. all filter outputs transmitted to a single electrode). Likewise, when the frequency allocation is shifted either apically or basally from its optimal frequency range (i.e. > 3 mm), there was a drop in speech recognition performance (Fu and Shannon, 1999). Simulations with normal hearing subjects have shown comparable results (Rosen et al., 1999).[1]

The second question concerns the degree of spectral resolution needed to represent an auditory image, and is related to the number of independent auditory information channels. Simulation studies demonstrate that normal hearing listeners presented with sentences in quiet achieve 90% accuracy with 4 frequency channels (sinusoids: Dorman et al, 1997; noise bands: Shannon et al, 1995) and 100% with 8 channels (Dorman et al, 1997). Likewise, in implant listeners the number of active electrodes appears to affect the level of speech understanding.[2] Several studies have shown that as the number of active electrodes was increased up to 4-7 electrodes, speech recognition scores improved (Fishman et al, 1997; Lawson et al, 1996). No further improvement was found as the number was increased from 7 to 20 active electrodes, even though the spectral information typically presented to all 20 electrodes was retained and transmitted to the reduced-electrode processors, i.e., no spectral information was discarded. Fishman et al. demonstrated that more electrodes were needed to reach a performance plateau with vowels (7 electrodes: 70-80% correct) than with consonants (2 electrodes: 95% correct) or sentences (4 electrodes: 69.3% correct). The best

[1] Since the frequency re-allocation studies included only postlinguistic adult patients, it is presently unclear whether the results were driven strictly by biological (or innate) constraints or were also, in part, due to a learned tonotopic representation. Early-implanted children with a frequency-to-electrode mismatch may differ biologically from adults and late-implanted children with the same degree of frequency shifting.

[2] With an increasing number of electrodes, there is a decrease in the width of the analysis filter, meaning that there is a narrower frequency range (i.e. enhanced frequency specificity) for each electrode

performance for implant listeners is equivalent to results with normal-hearing subjects presented with 1-7 frequency channels (Dorman et al, 1998a; Fu et al, 1998).

2.6 Multiple Electrodes and Speech Recognition

These laboratory-controlled studies appear to suggest that only a small number of channels would be sufficient to understand speech. However, even if the poorest implant performers could attain the equivalent of 4-7 frequency channels in laboratory tests, they would continue to experience difficulties in “real world” situations where listeners must accommodate to variations in fundamental frequency, speaking rate and pronunciation, along with variations in background noise levels. In fact, when normal hearing listeners were presented with an 8-channel simulation, their speech recognition scores for sentences dropped from 100% correct in quiet to 55% correct at a +2 dB signal-to-noise level (S/N), and only 16% correct at -2 dB S/N (Dorman et al, 1998). A minimum of 12 channels was required to achieve maximum performance for sentences in noise (Dorman et al, 1998), and greater than 16 channels for consonants and vowels in noise (Fu et al, 1998).

Loizou et al. (1999) examined the number of channels required for multitalker stimuli by presenting normal hearing listeners with 135 sentences, each with a different speaker. The mean speech recognition score with 4 channels was 63% correct, which was markedly lower than the score of 90% achieved with the single male talker used in previous studies. As mentioned in Section 2.4, each speech token, such as a vowel, produces a distinct pattern of spectral energy. With multiple talkers there is more variability in the formant pattern for a particular sound, and because of this variability, a small number of channels cannot adequately define a distinct pattern for each speech token. Taken together, these results indicate that implant patients should receive more than 4-7 channels of information as the degree of acoustic complexity increases.

One means of increasing the number of information channels is to simply increase the number of electrodes, since ideally cochlear implant listeners would have as many effective channels as they do electrodes. The Nucleus 24 multichannel cochlear implant, by Cochlear Corp., has 24 electrodes and the capability of stimulating up to 12 electrodes within a period of a few milliseconds. This implant has more electrodes than any other device currently available and should reasonably provide cochlear implant listeners with the greatest degree of spectral resolution. However, it was the Nucleus device that was used in the study by Fishman et al. (1997), which showed that speech recognition scores do not improve as the number of electrodes is increased beyond 4-7 electrodes. Fishman et al. discuss several possible reasons why this may be the case, i.e. ceiling effects, speech processor design limitations, trade-off between the number of electrodes and pulse rate, individual differences in pathology, and channel interaction. The rationale for suspecting electrical-field interaction is that as the number of electrodes was further increased beyond 7 electrodes, the probability for current field interactions between electrodes also increased, thereby negating any potential benefits afforded by additional electrodes.

2.7 Channel Interaction

With multichannel implants, selective stimulation is often imperfectly achieved even with the most modern cochlear implant designs and speech processing strategies. Both electrophysiological and psychophysical studies demonstrate that extensive current spread and channel interaction exists in multichannel prostheses (electrophysiological: Abbas, 1993; van den Honert and Stypulkowski, 1987; psychophysical: Dobelle, Brackmann, Mladejovsky, and Parkin, 1978; White, Merzenich, & Gardi, 1984). One type of channel interaction occurs when there is electric field summation in the cochlea. With simultaneous electrode stimulation, the current fields generated by the electrodes can add together and alter the loudness of the stimulus (White, Merzenich, and Gardi, 1984). Distortions in loudness might disrupt the listener's ability to encode spectral cues through comparisons of the relative amplitudes (or formant frequency energy) across electrodes. The second type of channel interaction can occur with non-simultaneous stimulation and is commonly referred to as neural-population or temporal interaction. White et al. (1984) demonstrated that substantial changes in threshold occurred when adjacent electrodes were stimulated within 2-5 ms of each other. The neurons were apparently responding as though they were receiving the pooled activity from each electrode. Therefore, when more than one electrode stimulates the same or overlapping neural populations, the stimuli may be perceptually indistinguishable or confused since the underlying neural populations would receive the combined stimulus and not the separate stimuli from each electrode (Leake et al, 2000; Shannon, 1993; Throckmorton and Collins, 1999; White and Van Compernolle, 1987; Zwolan, Collins, and Wakefield, 1997). This would limit the number of independent information channels conveyed to central auditory mechanisms and presumably impair the recognition of speech tokens coded predominantly by spectro-temporal cues, such as vowels and consonant place-of-articulation (Fishman et al., 1997; Fu et al, 1998).

Electrical-field interactions can occur even between non-adjacent electrodes if the current field emanating from a particular electrode is wide enough. In this case, the wide current field from one electrode may overlap, and therefore interact, with the current fields of its neighbors. The chance for electrical-field interactions decreases with interchannel distance, since the strength of the current field generated around an electrode decreases as it radiates away from the electrode (White et al, 1984). Current fields are generated between an active and a reference electrode. With bipolar electrode arrangements the reference and active electrode pair are both within the cochlea and are typically within a few millimeters of each other. In contrast, monopolar electrode arrangements typically have an active electrode within the cochlea and an extracochlear reference electrode surgically placed behind the ear in the temporalis muscle. The assumption is that a current field applied between adjacent bipolar electrode pairs will produce localized neural activation, whereas wider bipolar electrode separations will lead to broad current fields and a wider area of neural activation (Busby et al, 1994; van den Honert and Stypulkowski, 1987). Wider electrode separations between the active and reference electrodes reduce the current levels required to elicit an auditory sensation since the broader current fields have the potential to activate several nerve fibers. This effect has been demonstrated in both humans (Battmer et al., 1993; Lenhardt et al, 1992) and animals (van den Honert and Stypulkowski, 1987). With monopolar coupling, neural specificity is exchanged for lower current level requirements. In contrast, bipolar coupling provides the specificity but very few cells are activated. However, even a closely spaced electrode pair can activate a broad region of neural fibers at high current levels (van den Honert and Stypulkowski, 1987).

Current levels are relatively high when the distance from the electrode and target neural population is large (Shepherd et al., 1993).[3] In such cases, the current field must spread over a much wider range to reach the target cells. This problem may arise when the electrode array is positioned along the lateral wall of the cochlea during surgery or when there are only few surviving target neurons near the electrode. Wide current spread can potentially cause overlap between the current fields of simultaneously stimulated electrodes and in the neural populations excited by each electrode.

---

[3] High current levels are associated with wide current fields

2.8 Simultaneous vs. Sequential Speech Processing Strategies

The Continuous Interleaved Sampling (CIS) strategy presents sequential, non-overlapping pulses to each electrode (Wilson et al, 1991). The use of interleaved stimuli avoids the problem of electrical-field interaction common to speech strategies with simultaneous stimulation, such as the Compressed Analog (CA) speech strategy, since the biphasic pulses delivered to each electrode have temporal offsets that prevent vector summation of the electric fields. Comparisons between the CIS and CA speech strategy clearly demonstrate higher speech recognition scores with CIS even though the patients had only a few days of experience with the CIS strategy and 1-6 years of daily use with the CA strategy[4] (Boëx, et al, 1996; Schindler et al., 1995; Wilson et al., 1991).

The dilemma with the CA and CIS strategies is that monopolar coupling must be used to achieve adequate loudness growth. With a broad current field, it is not possible to accomplish selective stimulation of discrete neural populations, since several nerve fibers across the cochlea are recruited to achieve loudness. Interleaved stimulation also decreases the amount of speech information that can be delivered at a specific point in time. For instance, a very brief speech event occurring in a high frequency channel would not be transmitted to the listener if the electrode matched to that frequency is not being stimulated at that exact moment. Adequate loudness growth can be achieved for most patients with the "enhanced bipolar" coupling mode, used in the S-Series cochlear implant manufactured by Advanced Bionics Corp. With "enhanced bipolar" stimulation, the active medial electrode is paired diagonally with the lateral reference electrode (Fig 2.9.1).

[4] The CIS strategy also had a higher stimulation rate and a greater number of electrodes

The wider spacing between the active and reference electrode increases the breadth of the current field, but only slightly more than the field generated with the radial-bipolar orientation used in earlier CLARION® electrode designs, and markedly less than monopolar coupling. The increased number of neurons activated with "enhanced bipolar" stimulation, compared to the traditional radial-bipolar mode, reduces the amount of current required to reach threshold and achieve loudness growth. Consequently electrical-field interaction can be reduced for some patients.

A new speech processing strategy, called Simultaneous Analog Stimulation (SAS), was developed as a result of the new "enhanced bipolar" coupling method (see description in Appendix A). Approximately one-third to one half of the patients tested used SAS successfully (Battmer et al, 1999; Osberger et al, 1999). Higher levels of speech recognition were obtained for patients who preferred SAS compared to patients preferring the CIS strategy (Battmer et al, 1999; Osberger et al, 1999). Presumably the patients who used SAS had less electrical-field interaction than the patients who preferred CIS, although this was never directly tested. Figure 2.8 shows three speech processing strategies currently available for the Clarion cochlear implants.

Advanced Bionic's Commercial Speech Processing Strategies

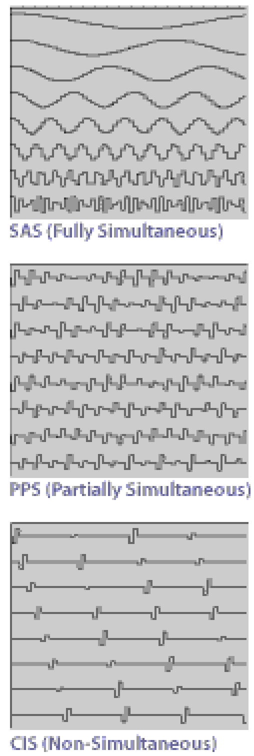


Figure 2.8: Speech processing strategies currently available with the Clarion cochlear implants. The degree of simultaneous stimulation ranges from purely simultaneous (SAS) in the top panel to purely sequential (CIS) in the bottom panel.

2.9 Reducing Electrical-field interaction through Electrode Design

Electrical-field interaction can also be reduced through cochlear implant electrode positioning and design. The goal is to place the electrode array closer to the modiolus and spiral ganglion cells so that less current is required to elicit an auditory percept (Rebscher et al., 1994). Reduced current spread increases the selectivity of electrical stimulation, and may increase the number of independent information channels and improve speech perception.

Advanced Bionics Corporation has recently developed an electrode "positioner" called the CLARION® Electrode Positioning System™ (EPS). The EPS consists of a "shim" inserted behind the electrode array that pushes the electrode array away from the outer wall of the scala tympani and closer to the habenula perforata (Figure 2.9.2).

Enhanced Bipolar Clarion Electrode

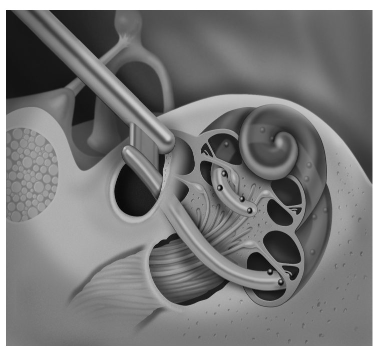


Figure 2.9.1: Schematic of the cochlea implanted with the standard Clarion electrode array. Enhanced bipolar coupling is achieved with the diagonally-arranged ball electrodes.

Enhanced Bipolar Clarion Electrode with the Electrode Positioner

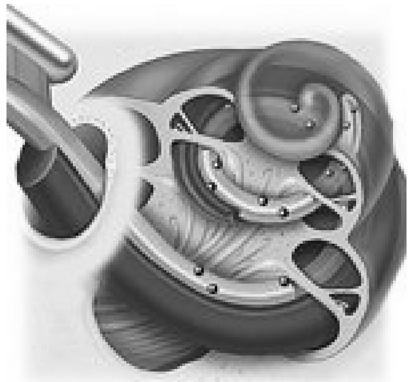


Figure 2.9.2: Schematic of the cochlea implanted with the standard Clarion electrode array (light blue) and the electrode positioner called the Electrode Positioning System (darker blue).

In addition to the EPS, new electrode geometries are being designed to focus the current field toward spiral ganglion cell bodies and prohibit current spread to neighboring electrodes (e.g. CLARION® Hi-Focus™). The Hi-Focus array has longitudinal plate electrodes separated by silastic buffers (Figure 2.9.3). The design and orientation of the electrodes directs most of the current towards the center of the modiolus. Current spreading to adjacent electrodes can be blocked by the silastic buffers, however this is only possible if the electrode array is resting against the modiolar wall. Therefore, the Hi-Focus array can only be effective when combined with the EPS.

Hi-Focus Electrode with the Electrode Positioner

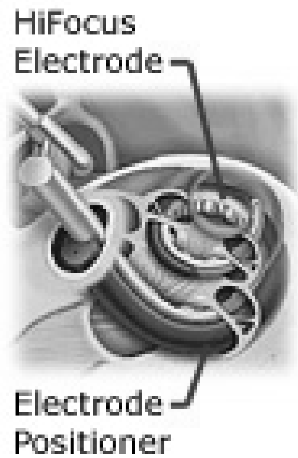


Figure 2.9.3: Schematic of the cochlea implanted with the Hi-Focus electrode array and Electrode Positioning System.

Clinical results with the Enhanced Bipolar electrode and EPS have shown lower stimulation levels and better speech perception scores than in patients without the EPS (Osberger et al, 1999). The clinical results also suggest that further benefit is achieved when the EPS is combined with the Hi-Focus electrode (Lenarz et al, 1999).

CHAPTER 3

Pilot Study

3.1 Introduction

A pilot study was conducted to examine the influence of electrical-field interaction on speech recognition performance. Electrical-field interaction was measured with a psychophysical task, known as simultaneous masking (Section 3.2). Psychophysical thresholds were compared for patients with the Hi-Focus electrode array and the Electrode Positioning System (HF+EPS), patients with the Enhanced Bipolar Clarion electrode and electrode positioner (ENH+EPS), and patients with the Enhanced Bipolar Clarion electrode without an electrode positioner (ENH). A correlation analysis was then carried out to test whether less electrical-field interaction was associated with higher speech recognition scores.

3.2. Simultaneous Masking

A psychophysical task, known as simultaneous masking, measures the degree of summation produced by two electrodes stimulated simultaneously. Several studies have used simultaneous masking to measure electrical-field interaction (Boëx et al., 1999; Shannon, 1983b, 1985; White, Merzenich, and Gardi, 1984). In a simultaneous masking paradigm, thresholds are compared for biphasic pulses, comprised of one anodic and one cathodic pulse, presented presented to a probe electrode alone and for pulses presented simultaneously to the probe electrode and a second electrode located some distance from the probe (i.e. the masker). The masker electrode either delivers biphasic pulses with the same phase first (Figure 3.2: middle panel) or is 180^{o} out-of-phase with the pulses of the probe electrode (Figure 3.2: bottom panel).

Simultaneous Masking Conditions

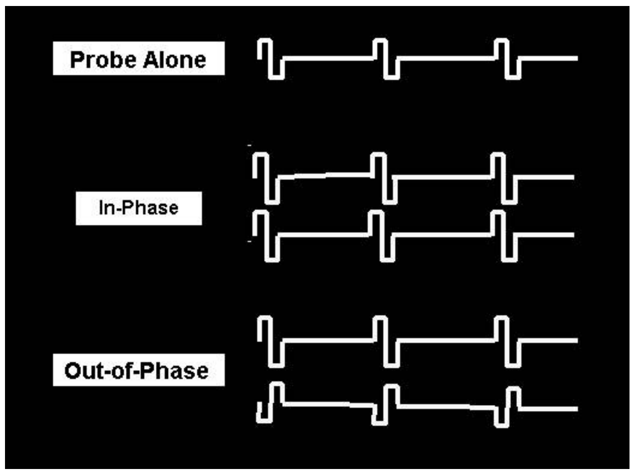


Figure 3.2: The top trace shows the biphasic pulse presented to a single electrode in the "probe alone" condition. The middle and lower traces show two pulses: one pulse is delivered to the probe electrode and the second is delivered to the masker electrode. The middle trace represents the "in-phase" condition, while the lower trace represents the "out-of-phase" condition.

If currents from the two electrical fields interact, then the listener's loudness percept is altered. For the out-of-phase condition, pulses would cancel and the percept becomes softer. In this case, the thresholds are elevated relative to thresholds for stimuli presented to the probe electrode stimulated alone. The opposite occurs for the in-phase condition. Specifically, when same-phase current pulses overlap, they add electrically, thereby increasing the loudness percept and lowering thresholds. The amount of electrical field overlap therefore determines how much current summation (in-phase) or cancellation (out-of-phase) occurs. The difference in thresholds for the in-phase and out-of-phase conditions represents the degree of current pulse overlap, or electrical-field interaction. The degree of electrical-field interaction can be expressed by the following formula:

$$\textit{Difference Threshold} = THR(-) - THR(+)$$

where,

THR(-) = out-of-phase threshold

THR(+) = in-phase threshold

In situations where there is virtually no electrical-field interaction, the threshold patterns would appear like those found with acoustical hearing (Shannon, 1983b, 1985). In acoustical hearing, when two stimuli activate two critical bands (or, similarly, two electrodes stimulating separate neural populations), the percept is louder and thresholds are lower than when only one critical band (or electrode) is activated. Phase does not affect the simultaneous masked thresholds in acoustical hearing. Both out-of-phase and in-phase stimuli decrease thresholds, and roughly to the same degree. Therefore, the *Difference Threshold* would approach zero as the amount of electrical field overlap decreased.

Simultaneous masking can be a useful measurement tool for evaluating the effectiveness of the EPS and Hi-Focus electrode array. If these new electrode designs serve their intended purpose, then it is predicted that the EPS combined with the Enhanced Bipolar Clarion electrode would generate less electrical-field interaction (or a smaller Difference Threshold) than the Enhanced Bipolar Clarion electrode without the EPS, and even less electrical-field interaction would occur when the EPS is combined with the Hi-Focus electrode array.

Many researchers have claimed that cochlear implant users with poorer speech recognition abilities may be subject to the detrimental effects of electrical-field interactions. Therefore, the relationship between electrical-field interaction and speech recognition performance was also evaluated in the pilot study. It was hypothesized that cochlear implant users with more electrical-field interaction would show lower speech perception scores than users with less electrical-field interaction.

3.3. Methods

Simultaneous masking measures were collected from 5 subjects implanted with the Enhanced Bipolar Clarion electrode without the electrode positioning system (ENH), 5 subjects with the Enhanced Bipolar Clarion electrode with the electrode positioning system (ENH+EPS), and 4 subjects with the Hi-Focus electrode with the electrode positioning system (HF+EPS). Subject CS of the HF+EPS group has not returned to complete the vowel and consonant portion of the experiment. Subject demographics are shown in Table 3.3.

Subject	Age	Electrode Type	Speech Strategy	Duration of HL (yrs)	Duration of Deafness (yrs)	Duration of CI Use (yrs)
JW	51	ENH	CIS	8	8	1.2
KH	54	ENH	CIS	5	5	1.1
EC	65	ENH	CIS	.2	.2	.5
VC	54	ENH	CIS	18	8	1
MK	58	ENH	CIS	7	7	2.3
MS	46	ENH+EPS	CIS	2	2	.7
MI	46	ENH+EPS	CIS	20	10	.7
SM	42	ENH+EPS	CIS	25	25	.6
SL	48	ENH+EPS	CIS	1	1	.7
BH	59	ENH+EPS	CIS	1	1	.5
BD	56	HF+EPS	PPS	39	8	.4
CS	37	HF+EPS	PPS	7	7	.2
SH	23	HF+EPS	SAS	.7	.7	.6
JPM	31	HF+EPS	CIS	1	1	.9

Table 3.3: Pilot Subject Demographics

The magnitude of electrical-field interaction was determined by measuring the amount of simultaneous masking between *adjacent* electrodes. Measuring electrical interactions between adjacent electrodes represents the extreme case, since electrical-field interaction typically decreases with increasing masker probe separations. Three masker electrode locations were tested: electrode 2 (apical masker), electrode 4 (middle masker), and electrode 6 (basal masker). The corresponding probe electrodes were as follows: probe electrodes 1 or 3 paired with masker electrode 2; probe electrodes 3 or 5 paired with masker electrode 4; and probe electrodes 5 or 7 paired with masker electrode 6. This produced six masker+probe conditions. Simultaneous masking tasks were completed for monopolar and bipolar stimulation modes. This resulted in a total of twelve conditions (6 masker, probe pairs x 2 stimulation modes).

Several speech perception measures in quiet and in noise were included to explore the relationship between channel interaction and various speech cues. Percent correct scores were obtained for HINT sentences (Nilsson et al., 1994), CNC words, vowels (/hVd/) taken from the materials collected by Hillenbrand et al. (1995), and Iowa consonants in /aCa/ environment taken from a set developed by Shannon et al (1999). Speech testing was performed in a sound-attenuated chamber. Speech stimuli were delivered through loudspeakers at a 0^{o} azimuth and presented at 65 dB(A) in quiet and in noise at +5 and +10 dB signal-to-noise ratios. All speech testing was performed with the patient's own speech processing strategy.

3.4 Results and Discussion

3.4.1 Electrical-Field Interaction Results

In Figures 3.4.1.1 and 3.4.1.2, the magnitude of electrical-field interaction, as measured by the *Difference Threshold*, is plotted for each electrode design and each masker electrode location (e.g. apical masker = electrode 2, middle masker = electrode 4, basal masker = electrode 6). In these two figures, the magnitude of electrical-field interaction for each masker electrode location was obtained by averaging the *Difference Threshold* for the two masker+probe pairs with the same masker electrode. For example, the *Difference Threshold* value for the apical masker location was obtained by averaging the *Difference Threshold* for probe electrode 1 with masker electrode 2 and probe electrode 3 with masker electrode 2. Figure 3.4.1 shows the electrical-field interaction data for the monopolar design and Figure 3.4.2 shows the data for the bipolar design.

These results indicate that, regardless of the electrode design, the HF+EPS subjects had the least electrical-field interaction, and the ENH group had the most electrical-field interaction. The results also show that the broad current distribution of monopolar stimulation produces relatively uniform levels of electrical-field interaction for each of the masker electrodes. With monopolar stimulation, a large portion of the modiolus lies within the current field and, because of this, the nerve fibers along the entire length of the cochlea could be activated. Therefore, regardless of the masker+probe pair, the same nerve populations contribute to the threshold measures in the simultaneous masking task. In contrast, with bipolar stimulation, the degree of electrical-field interaction varied depending on the location of the masker. Since only a small region of auditory nerve fibers are activated with bipolar coupling, there was more variability in the magnitude of channel interaction for different masker+probe pairs.

Pilot Study: Monopolar Electrical-Field Interaction

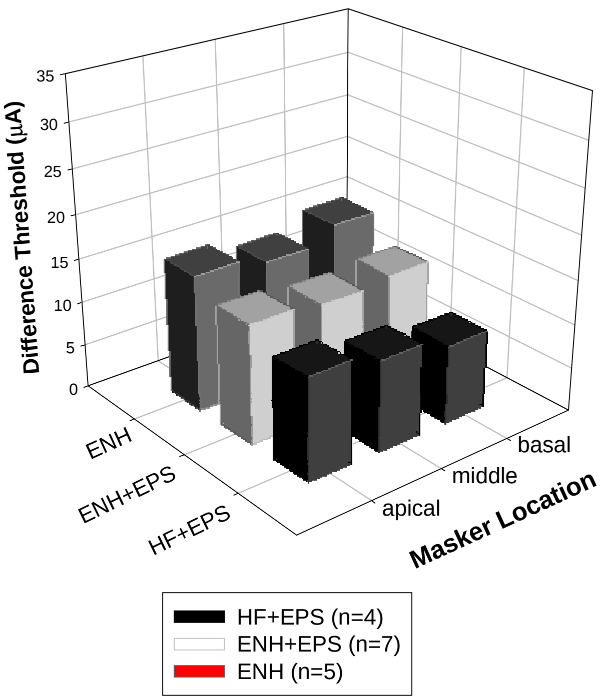


Figure 3.4.1.1: The magnitude of **monopolar** electrical-field interaction (represented by the Difference Threshold) is shown for each of the three electrode designs and the three masker locations.

Pilot Study: Bipolar Electrical-Field Interaction

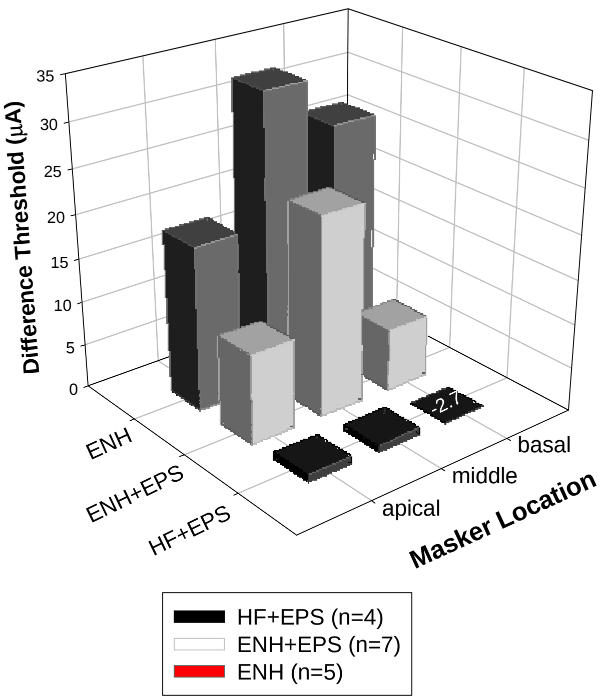


Figure 3.4.1.2: The magnitude of **bipolar** electrical-field interaction (represented by the Difference Threshold) is shown for each of the three electrode designs and the three masker locations.

3.4.2 Speech Recognition Results

Figure 3.4.2 shows speech recognition performance in quiet and in noise for vowels, consonants, and sentences. These data were collected from subjects who were regular users of their device for at least 6 months. Since the HF+EPS subjects did not have sufficient experience with their device at the time of testing, data is shown for 5 ENH and 4 ENH+EPS subjects only. Speech recognition testing was performed with the monopolar CIS speech processing strategy, since all subjects were regular users of this strategy. The results demonstrate that higher speech recognition scores were obtained for the ENH+EPS group compared to the ENH group. This result was clearly apparent for all but the “vowels in quiet” condition, presumably because monopolar stimulation is not as sensitive to spectral distortions as bipolar stimulation. In addition, the ENH+EPS group maintained high sentence recognition scores with the addition of speech-shaped noise, while the ENH group showed a large drop in performance.

Pilot Data: Speech Recognition Performance

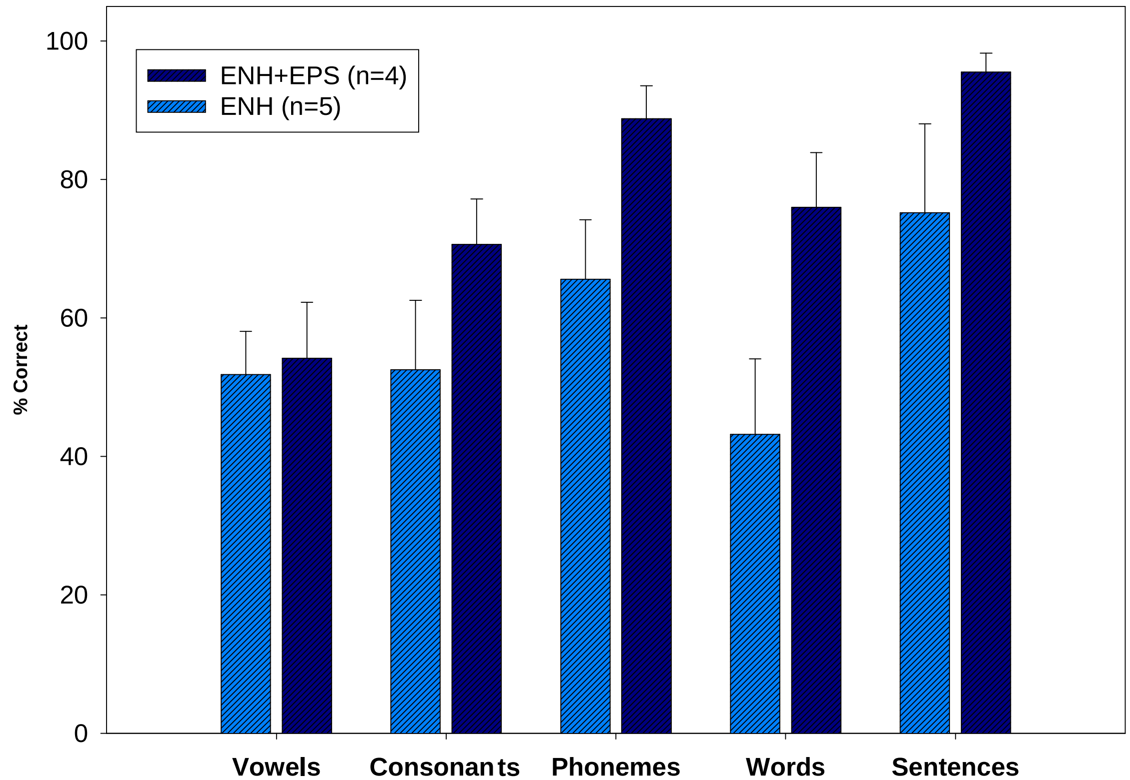


Figure 3.4.2: Mean speech recognition performance in quiet is shown for the five speech recognition tasks. Results are shown for four patients with the enhanced bipolar Clarion electrode array with an electrode positioner and the five patients without an electrode positioner. Error bars show the standard error for each group and speech task.

3.5 Relationship Between Electrode Interaction and Speech Recognition

Since all speech testing was performed with the monopolar CIS strategy, correlations were made between the monopolar electrode interaction data and speech recognition scores. Figure 3.5 shows a scatterplot of the correlations between speech recognition performance and channel interaction. Moderate to strong negative correlations were found between the degree of electrode interaction and speech performance. The strongest correlations were found for consonants and sentences. These results indicate that as electrical-field interaction decreases, speech recognition performance increases. Overall, the ENH+EPS subjects tended to have higher speech scores and less electrode interaction than the ENH subjects.

Pilot Study: Speech Recognition as a Function of Electrical-Field Interaction

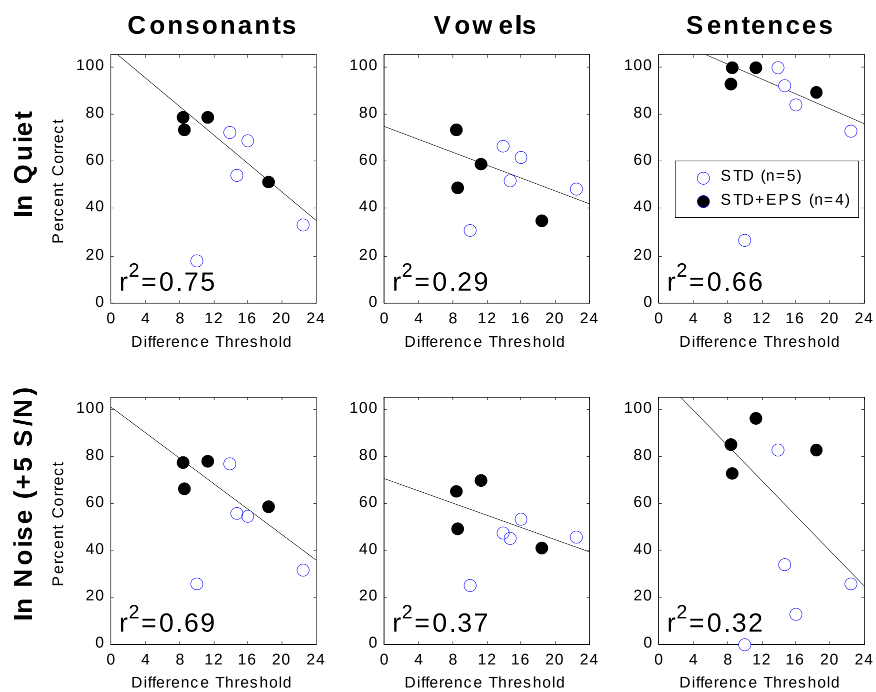


Figure 3.5: Scatterplots of the correlations are shown for the speech recognition tasks tested in quiet (top) and noise (bottom). The average difference threshold across all monopolar simultaneous masked conditions is plotted as a function of the percent of items correctly identified for each speech task with a monopolar CIS speech strategy. Filled circles represent the data for each of the four ENH+EPS subjects and unfilled circles represent the data for each of the four ENH subjects.

3.6 Summary

These preliminary experiments have examined the effects of electrical-field interaction on speech recognition. Speech recognition scores were strongly correlated with the magnitude of electrical-field interaction between *adjacent* electrodes. It was also shown that the magnitude of electrical-field interaction varies as a function of electrode design. Electrical-field interaction was greatest for patients with the ENH electrode design and least for patients with the HF+EPS electrode design.

In the experiments that follow, the extent of electrical-field interaction across the electrode array, instead of the magnitude of electrical-field interaction between adjacent electrodes, is investigated for the three electrode designs. The relationship between electrical-field interaction and speech processing strategy performance is also evaluated. Subjects were tested with 5 speech processing strategies varying in the number of simultaneous channels. Since simultaneous stimulation increases the likelihood for electrical-field interactions to occur, it was predicted that a high degree of electrical-field interaction should produce higher speech recognition scores for more sequential as opposed to simultaneous speech strategies. Specifically, it was predicted that subjects with lower levels of electrical-field interaction would be able to take advantage of the greater spectral and temporal resolution provided by simultaneous speech strategies, and would therefore show higher speech scores for simultaneous than sequential speech strategies.

CHAPTER 4

MEASURING ELECTRICAL-FIELD INTERACTIONS WITH SIMULTANEOUS MASKING

4.1 Introduction

Minimizing electrical-field interactions in a multichannel cochlear implant is one of the potential advantages of spatially restricting the electrical field potential. When selective electrode stimulation is not achieved, the perceptual response to one electrode can affect the perceptual response to stimulation applied to other electrodes (Collins et al., 1997; Cotter, 1986; Eddington et al., 1978; Nelson et al., 1995; Throckmorton and Collins, 1999; Tong et al., 1983). For example, if the electrode array emits extensive current spread, then the current fields from two or more simultaneously stimulated electrodes will overlap and summate, affecting the perception of loudness, pitch, and possibly even the quality of speech (Shannon, 1993). Electrical-field interactions can therefore produce current flow patterns distinct from those arising from each of the single, isolated electrodes and as a consequence, auditory nerve fiber activity in the region of overlap may not be similar to the activity produced by electrodes stimulated individually.

In this chapter, as with the pilot study, a simultaneous masking task was used to examine electrical-field interactions in patients with the CLARION multichannel cochlear implant (refer to Section 3.1). However, instead of measuring the magnitude of electrical-field interaction between adjacent electrodes, electrical-field interaction was measured as a function of the masker+probe separation. Based on previous research electrical-field interaction should decrease as the separation between simultaneously stimulated electrodes increases (White et al, 1984). The electrode separation at which electrical-field interactions dropped to negligible levels was of particular interest in this study. The "sharpness" of this function was expected to vary across patients, e.g. some patients might have electrical-field interactions only between adjacent electrodes while others could experience interactions even with the maximum masker+probe separation.

In Section 2.9, it was argued that the design of the electrode array is crucial for reducing electrical-field interaction. In this experiment, it was hypothesized that the extent of electrical-field interaction for patients with an electrode positioner (EPS) would be less than for patients without the EPS. This is because the primary purpose of the positioner is to decrease the distance between the electrode array and target cells within the modiolus. This electrode modification should reduce the amount of current required to elicit an auditory sensation and limit current spread across the electrode array, thereby constraining electrical-field interaction to probe electrodes closest to the masker. The design of the Hi-Focus electrode, when combined with the EPS, is expected to further reduce the extent of electrical-field interaction. Theoretically, since silastic buffers are placed between adjacent electrodes in the Hi-Focus array, current should be constrained and unable to interact with the current field from its neighbor. For this reason, electrical-field interaction should be at negligible levels even between adjacent electrodes.

The extent of electrical-field interaction will, however, vary as a function of the coupling mode used in the simultaneous masking task. Since monopolar stimulation produces a broad current field between the active electrode within the scala tympani and the reference electrode outside the cochlea, current is applied over a substantially greater area than with bipolar stimulation. It is predicted that electrical-field interaction will exist across the entire electrode array with monopolar coupling, regardless of the electrode design. In contrast, bipolar stimulation should show large differences between the three electrode designs because of the unique bipolar electrode arrangements used in each design and the current fields they produce.

4.2 Method

4.2.1 Subjects

Thirteen postlingually-deafened adults (20-69 years of age) were selected for these experiments. Each electrode design group had 3 users of either the CIS or PPS speech strategy. Three ENH SAS users and one HF+EPS SAS user were also included to compare with the results from the CIS and PPS users (Section 4.3.3). Patients were recruited from the House Ear Clinic in Los Angeles, CA and from the Callier Center for Communication Disorders in Dallas, TX. All subjects were native English speakers with at least 5-months experience with their device. An informed consent was obtained and participants were paid on an hourly basis. Patient demographics are shown in Table 4.2.1.

4.2.2 Stimuli and Equipment

A simultaneous masking task was used to measure the extent of electrical-field interaction across the electrode array (see Section 4.1.1). In this task, a series of charge-balanced, biphasic pulses (300 μsec/phase, 200 msec burst duration, and a 1000 Hz rate) were presented to the electrodes, using either monopolar or bipolar stimulation. The stimuli were driven via custom-designed software developed by Advanced Bionics Corp., called the Electrode Interaction Tester (EIT). The EIT software runs on a personal computer and interfaces with the patient's S-Series speech processor through the Clarion Processor Interface (CPI). Subjects were seated in a soundproof chamber for the duration of the task.

4.2.3 Procedure

Prior to testing, thresholds were obtained for each of the electrodes, separately for monopolar and bipolar stimulation using an adaptive clinical procedure, commonly referred to as the modified Hussen-Westlake technique (Carhart and Jerger, 1959). This procedure estimates the amount of stimulation capable of evoking a response 50% of the time. The pulse amplitude delivered to the masker electrode (i.e. electrode 4) was then fixed at 70% of its threshold, while the amplitudes of the remaining probe electrodes (i.e. electrodes 1, 2, 3, 5, 6, 7) were initially set to 10 clinical units (cu) above their respective thresholds. A clinical unit is simply a monotonic, logarithmic increment in current amplitude or pulse width that is roughly related to dB level and is patient-dependent.

In the test session, a 3-interval forced choice adaptive tracking procedure was used to obtain thresholds for pulses presented to the probe electrode alone and for pulses presented to the probe and masker electrodes simultaneously. In the simultaneous condition, pulses delivered to the masker electrode were either “in-phase” or “out-of-phase” with the probe electrode (see Section 3.1). Therefore, there were three simultaneous masking conditions (Probe Alone, In-Phase, and Out-of-Phase) for each stimulation mode (monopolar and bipolar) and for each probe electrode (electrodes 1, 2, 3, 5, 6, 7). All conditions were randomized. Subjects responded by pressing a mouse button to indicate which interval contained the stimulus. Two consecutive correct decisions led to a decrease in the probe electrode’s pulse amplitude (or loudness) and one error increased pulse amplitude. Visual feedback was provided after each trial. This procedure estimated the amount of stimulation current required for 70.7% correct responses (Levitt, 1971). The last 8 reversals were averaged to compute the threshold for each condition.

Subject	Age	Electrode Type	Speech Strategy	Duration of HL (yrs)	Duration of Deafness (yrs)	Duration of CI Use (yrs)
DF	66	ENH	PPS	43	5	2.2
EC	67	ENH	CIS	.2	.2	2
VC	55	ENH	CIS	18	8	2.1
VR	62	ENH	SAS	20	6	2.3
MC	43	ENH	SAS	.2	.2	2.5
GR	44	ENH	SAS	20	2	4
CSM	57	ENH+EPS	CIS	35	18	1.4
RGL	46	ENH+EPS	CIS	.2	.2	.4
MI	47	ENH+EPS	CIS	20	10	1.7
BC	49	HF+EPS	CIS	46	29	.7
ML	68	HF+EPS	PPS	34	26	.4
BD	57	HF+EPS	PPS	39	8	1.1
PD	59	HF+EPS	SAS	36	36	1.2

Table 4.2.1: Patient Demographics

4.3 Results and Discussion

4.3.1 Simultaneous Masked Thresholds

Averaged thresholds across the nine CIS users and six masker+probe pairs are shown in Figure 4.3.1 (left panel: monopolar thresholds; right panel: bipolar thresholds). The mean bipolar thresholds were between 15 to 20 μA higher than monopolar thresholds. This difference in current amplitude is related to the width of the current field generated by these two types of stimulation. As mentioned previously, bipolar current fields flow between closely spaced, intracochlear active and reference electrodes. The closeness of the active and reference electrodes with bipolar coupling restricts the current field to a very limited region of the cochlea. In contrast, monopolar current fields spread between an active electrode within the cochlea and a remote reference electrode outside the cochlea, producing a broad current field. Because of the relatively extensive area of neural activation with a monopolar current source, less current is generally required to elicit a sensation of sound. Also of interest is the difference in the error bars between the two types of stimulation. When the excitation pattern is broad, the effects of local neural survival tend to average out, resulting in less threshold variation between subjects. Bipolar stimulation is more sensitive to these anatomical differences and consequently produces larger error bars.

More importantly, monopolar and bipolar stimulation yield unique threshold patterns across the three masking conditions. Threshold patterns indicative of electrical-field interaction would be lowest for the in-phase condition and highest for the out-of-phase condition due to the summation of current with overlapping current fields (see Section 2.1). This pattern is clearly demonstrated with the pooled monopolar data in the left panel of the figure. A different pattern emerges with the pooled bipolar data, shown in the right panel. Thresholds for the bipolar In-Phase and Out-of-Phase conditions are relatively the same, and lower than for the Probe Alone condition. This particular pattern represents the case for no electrical-field interaction. Without current field overlap, the threshold is independent from the polarity of the current and the Out-of-Phase and In-Phase conditions produce similar thresholds.

Simultaneous Masked Thresholds

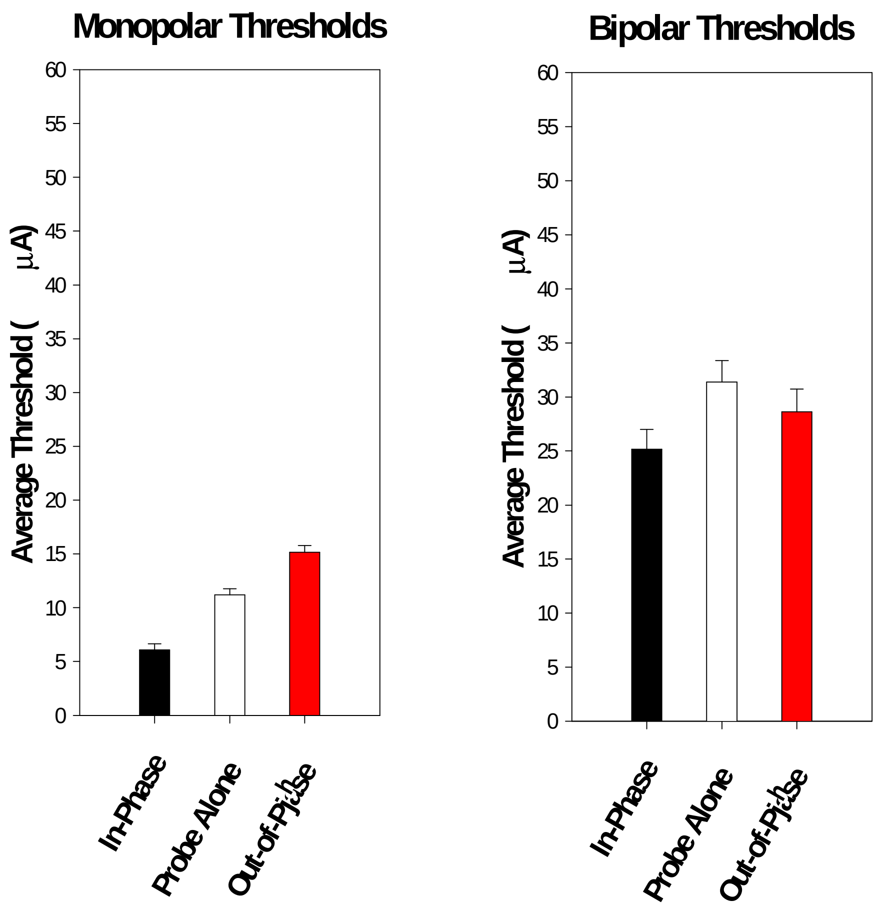


Figure 4.3.1.1: The threshold data in these figures is the average threshold across all six masker+probe pairs and all nine CIS users. The In-Phase condition represents the case where the masker and probe electrodes generate current fields with the *same* current field polarity. The Out-of-Phase condition is when the masker and probe have the *opposite* polarity. For the Probe Alone condition, only the probe electrode is stimulated. Error bars show the standard error across subjects and the six masker+probe pairs.

The simultaneous masked thresholds are shown separately for each electrode design in Figure 4.3.1.2 (monopolar) and Figure 4.3.1.3 (bipolar). Thresholds shown in the left panel of the figure are for the HF+EPS group, the middle panel shows the results for the ENH+EPS group, and the panel on the right shows thresholds for the ENH group. With monopolar stimulation, all the electrode types had thresholds indicative of electrical-field interaction (e.g. highest thresholds for the Out-of-Phase condition and the lowest thresholds for the In-Phase condition). With bipolar stimulation, each electrode design had a slightly different pattern. The difference in thresholds (i.e. *Difference Threshold*) between the Out-of-Phase and In-Phase condition were smallest for the HF+EPS group (-1.16 μA) and largest for the ENH group (7.95 μA ENH). The *Difference Threshold* for the ENH+EPS group was 3.64 μA. As mentioned in Section 2.1, the size of the *Difference Threshold* is related to the magnitude of electrical-field interaction. The results therefore suggest that patients with the Hi-Focus electrode design had less electrical-field interaction than either of the patient groups with the Enhanced Bipolar design. In addition, patients with the electrode positioner (or EPS) had less electrical-field interaction than patients without the EPS.

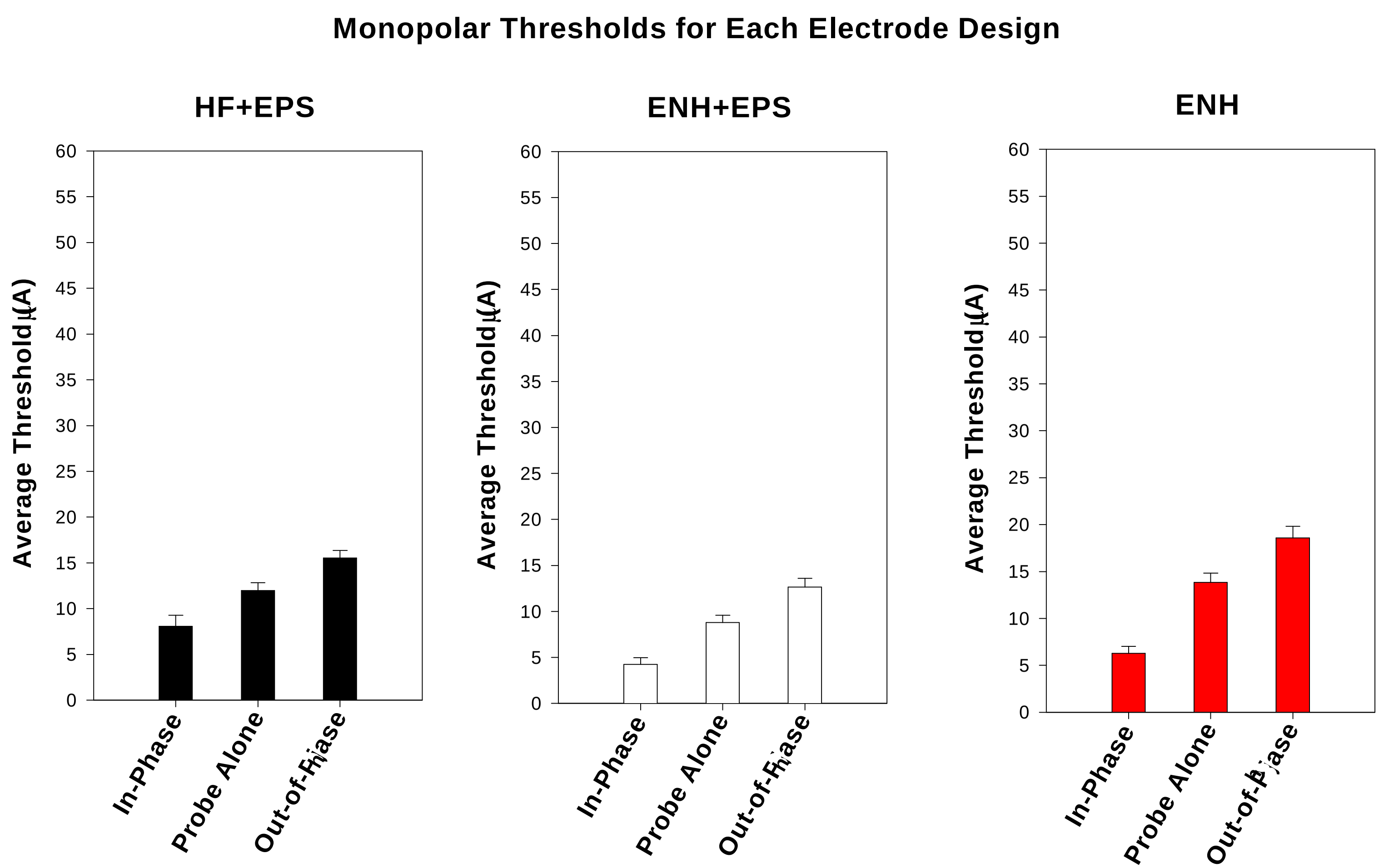


Figure 4.3.1.2: This figure shows the monopolar simultaneous masked thresholds for each of the three electrode designs. Error bars represent the standard error across the three subjects and six masker+probe pairs.

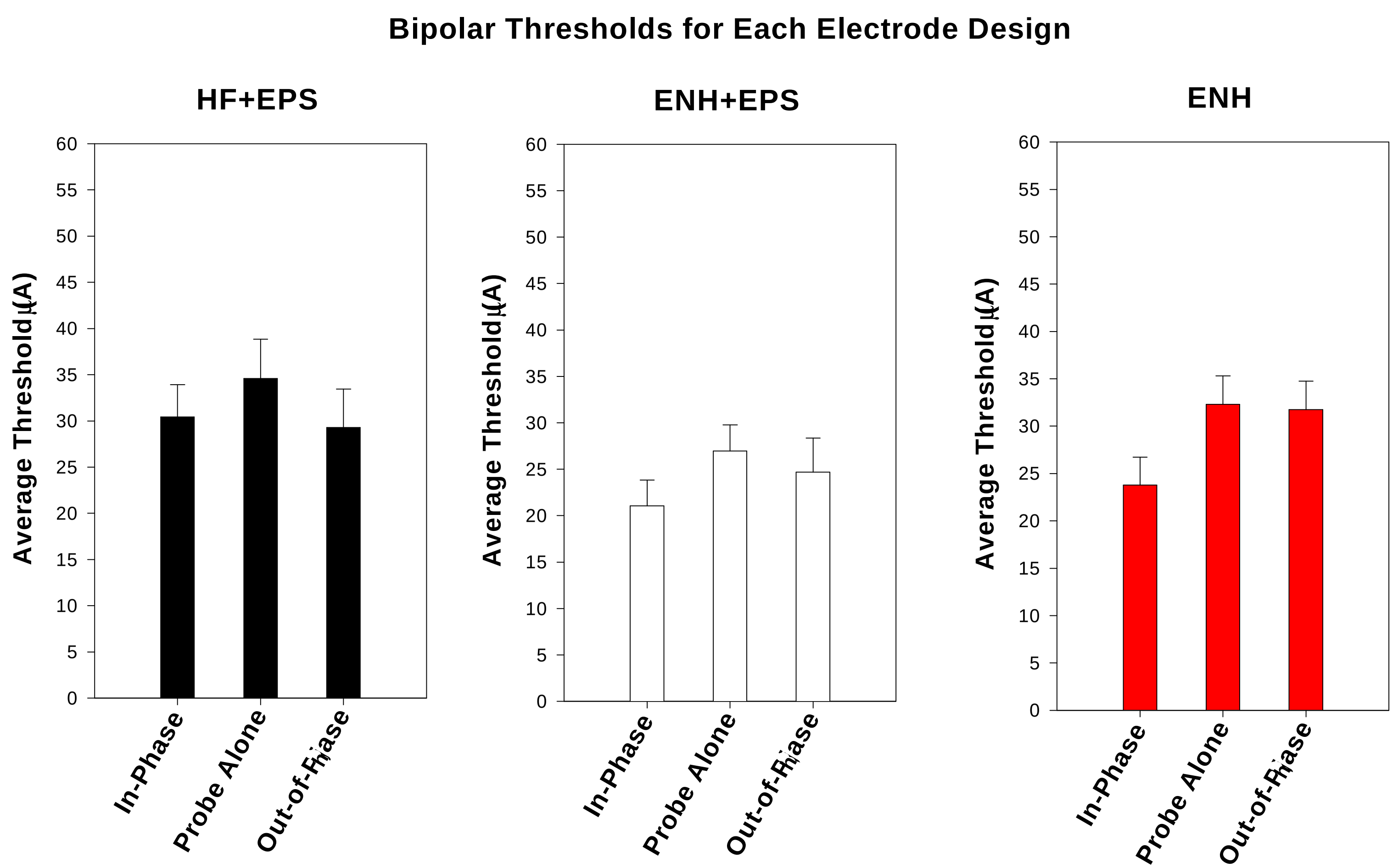


Figure 4.3.1.3: This figure shows the bipolar simultaneous masked thresholds for each of the three electrode designs.

4.3.2 Pattern of Interaction Spread

Electrical-field interaction spread is the change in electrical-field interaction magnitude (measured with the simultaneous masked *Difference Threshold* = Out-of-Phase - In-Phase) as a function of the distance between the probe and masker electrode. Figure 4.3.2.1 compares electrical-field interaction spread for each of the three electrode designs. Figure 4.3.2.2 shows the same electrical-field interaction data from Figure 4.3.2.1, but normalized according to the following interaction index formula ([need more refs] Boex et al., 1999):

$$\mathbf{II = (T_{P(-)} - T_{P(+)}) / 2*ML}$$

Where,

II: Interaction Index

$T_{P(-)}$: Out-of-Phase Probe Threshold

$T_{P(+)}$: In-Phase Probe Threshold

$T_{P(-)} - T_{P(+)}$: Difference Threshold

C_m: Masker Current Level

The derivation of this formula is as follows:

(1) When the electrical fields completely interact,

$T_{P(-)} = C_p + C_m$, where C_p is the Probe Current Level

$T_{P(+)} = C_p - C_m$

And,

$II = (C_p + C_m) - (C_p - C_m) / (C_p + C_m) - (C_p - C_m)$

$= 2* C_m / 2* C_m$

$= 1$

(2) When the electrical fields are independent,

$T_{P(-)} = C_p + C_m$

$T_{P(+)} = C_p + C_m$

And,

$II = (C_p + C_m) – (C_p + C_m) / (C_p + C_m) – (C_p + C_m)$

$= 0$

The pattern of results remained unchanged after normalizing.

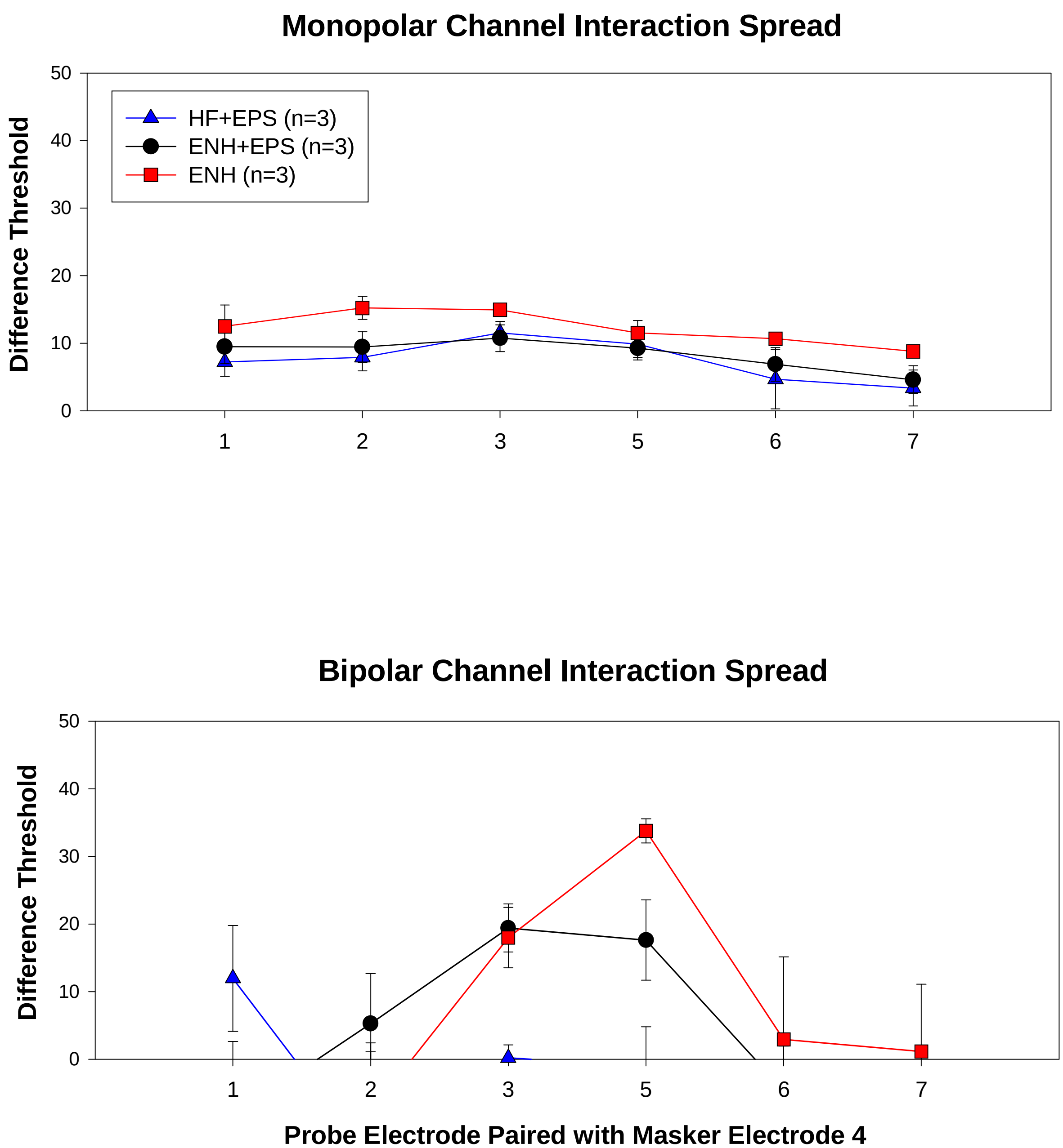
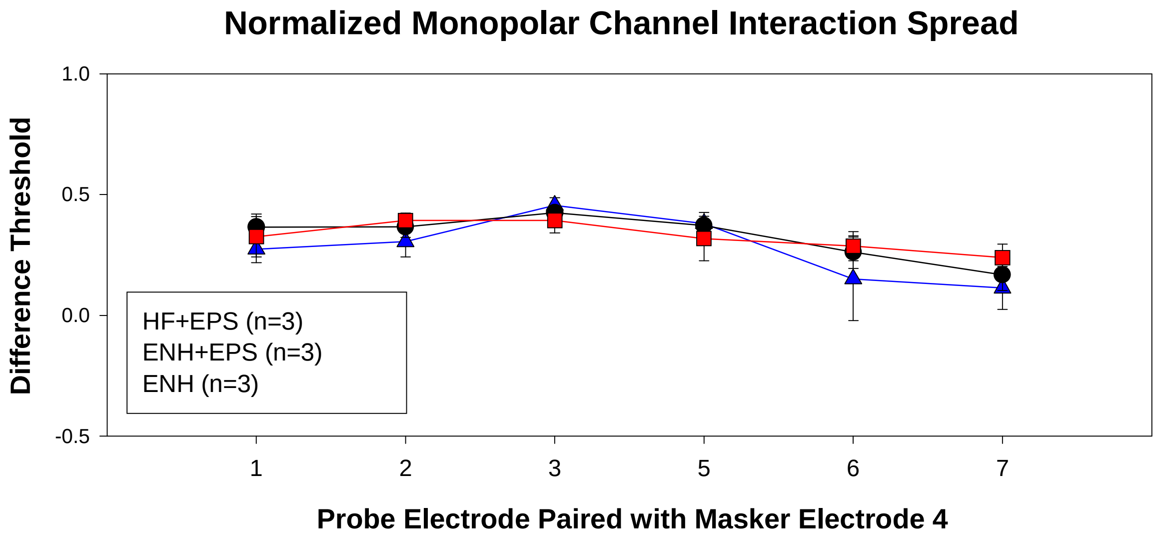


Figure 4.3.2.1: Monopolar interaction spread is shown in the top panel and bipolar interaction spread is shown in the bottom panel. The y-axis is the magnitude of channel interaction given by the formula for the masked difference threshold. The x-axis is the probe electrode, which was paired with masker electrode 4 located in the center of the electrode array. The magnitude of channel interaction decreases as the separation between the masker and probe electrode increases.



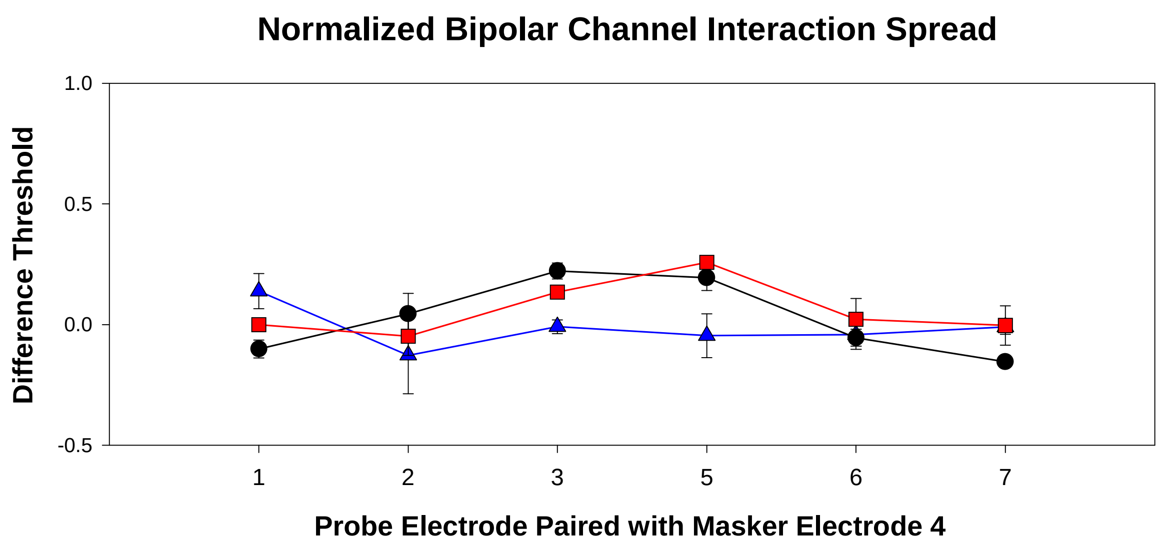


Figure 4.3.2.2: This figure shows the normalized data from Figure 3.3.2.1

A 3x2 mixed factorial design (Electrode **Design**: HF+EPS, ENH+EPS, and ENH; Stimulation **Mode**: bipolar and monopolar) was used to analyze the data. The dependent variable was the **area (in μA x mm)** under the interaction spread curve, calculated separately for each subject. Negative electrical-field interaction magnitudes, such as those found for most of the HF+EPS subjects, were set to zero prior to making the area calculation. There were significant main effects of design [$F(2,6)=6.01$, $p<0.05$] and mode [$F(1,6)=69.79$, $p<0.001$] with no significant interaction of design by mode [$F(2,6)=0.03$, $p>0.05$].

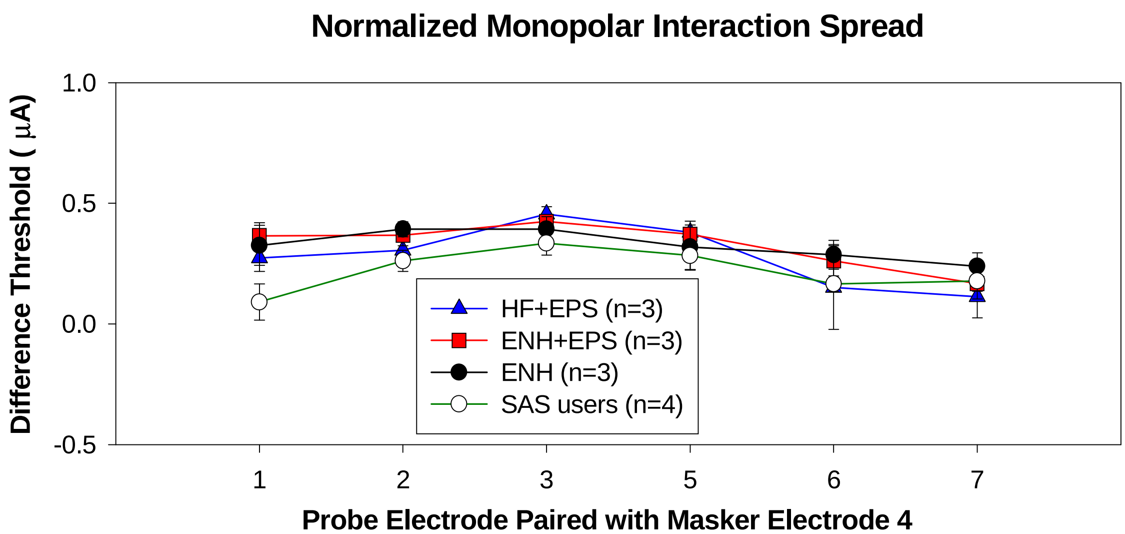
In agreement with previous findings, the magnitude of electrical-field interaction decreased with greater masker, probe separations (Shannon 1983, 1985; White et al., 1984). This was most apparent with the bipolar coupling mode. The narrow current field of bipolar electrode pairs was less likely to produce interactions between distant electrodes than current from monopolar electrode pairs. As such, the pattern of interaction with bipolar coupling showed a more abrupt decrease in interaction magnitude with increasing electrode separation compared to the same case with monopolar coupling.

Planned comparisons demonstrated that the HF+EPS group had less interaction spread than those with the ENH electrode design regardless of the stimulation mode [$F(1,6)=12.01$, $p<.05$]. In the case of bipolar stimulation, the HF+EPS group had negligible levels of interaction even between *adjacent* masker+probe pairs. This indicates that, at threshold, electrical-field interaction with Hi-Focus design was substantially reduced or absent for electrode separations greater than 1.1 mm from the probe electrode (Note: 1.1 mm is the separation between adjacent electrodes in the Hi-Focus design). In contrast, the ENH and ENH+EPS groups required masker+probe separations of at least *two* electrodes for electrical-field interaction to drop to negligible levels. This is equivalent to an electrode separation of at least 3 mm (Note: there is a distance of 1.5 mm between adjacent active electrodes in the Enhanced Bipolar design). This potentially important preliminary result suggests that, in the Hi-Focus electrode array, the use of silastic buffers between adjacent electrodes reduced electrical-field interaction considerably.

4.3.3 Electrical-Field Interaction: Differences Between CIS and SAS Users

The normalized interaction spread data, with the SAS users included, is shown in Figure 4.3.3. The SAS users in this study had less electrical-field interaction across the electrode array than most of the CIS/PPS users. This result is consistent with the claim that only patients with low levels of electrical-field interaction would prefer a simultaneous speech processing strategy. It is also important to point out that the subjects from the HF+EPS group had comparable interaction levels. However, these patients preferred and now regularly use the PPS strategy, not SAS! This apparent contradictory result may be because SAS and CIS were the only two speech processing strategies initially available to patients with the Enhanced Bipolar electrode. SAS and CIS are from opposite ends of the speech processing strategy scale - SAS is a fully simultaneous strategy and CIS is sequential. Patients with the Enhanced Bipolar electrode who had lower levels of interaction probably chose SAS even though it was not the best match. When the Hi-Focus electrode was introduced, PPS was already on the market. PPS is partially simultaneous, where two, non-adjacent electrodes are stimulated at the same time. Perhaps if PPS were available to the ENH SAS users at the time of initial implant stimulation, they would now be PPS and not SAS users.

Interaction Spread for SAS Users



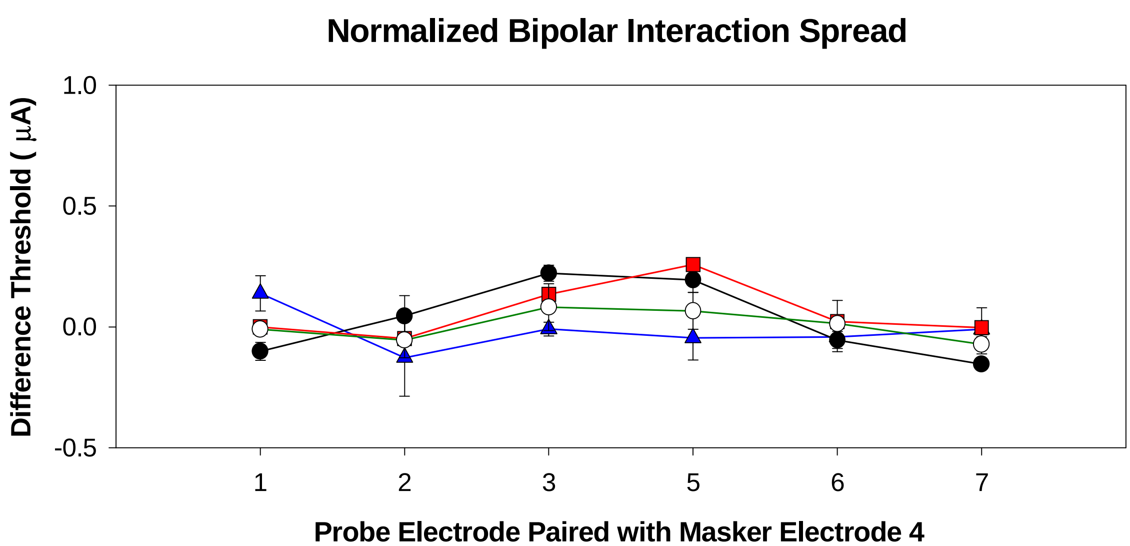


Figure 4.3.3: Comparison of interaction spread for SAS (green line with open circle) and CIS users.

4.3.4 Comparisons with Pilot Data

The pilot study assessed the magnitude of electrical-field interactions occurring between adjacent electrodes at several points along the cochlea, whereas the current study investigated the electrode separation necessary for electrical-field interactions to drop to negligible levels. Both studies demonstrate that bipolar electrical-field interaction for the HF+EPS group was either very low or absent between adjacent electrodes. In contrast, the ENH and ENH+EPS groups required at least a 3 mm electrode separation to reach comparable levels. These two studies support the hypothesis that patients with the Hi-Focus electrode will have lower levels of electrical-field interaction, and therefore less interaction spread, than patients with the Enhanced Bipolar electrode.

The pilot data suggested that patients with the electrode positioner (EPS) have lower levels of electrical-field interaction than patients without the EPS, however this result was not upheld in the present study. Patient-specific variables might have obscured the effects of the positioner in the present study. Two of the three new ENH+EPS subjects were “poorer” users than the previous ENH+EPS subjects (n=5). Subject CS from the present study had difficulty maintaining lock between the external headpiece and the implant receiver due to a thick skin flap. Patients with headpiece retention problems often report that the sound fluctuates between being muffled and clear. This problem most likely resulted in higher thresholds in the simultaneous masking, elevating the overall thresholds for the ENH+EPS users compared to those in the pilot study.

4.3.5 Future Directions

The simultaneous masked thresholds were useful for assessing the magnitude and extent of electrical-field interactions at low stimulation levels and for evaluating electrode modifications. However, to obtain a better understanding of the electrical-field interactions encountered with conversational speech levels, a different measurement tool is needed. This is because Advanced Bionics has built in a safety feature that limits the amount of current that can be delivered to the electrodes. The Out-of-Phase condition at higher masker and/or probe levels might reach this plateau, making it difficult to establish a measurement.

Advanced Bionics has recently introduced the second generation Research Interface (CRI II). The CRI II includes tools that can be used to make safe measurements of electrical-field interactions at higher current levels. Ideally, this new device could be used to model electrical-field interaction at several levels within the patient's dynamic speech range.

CHAPTER 5

THE INFLUENCE OF SPEECH PROCESSING STRATEGIES ON SPEECH RECOGNITION PERFORMANCE

5.1 Introduction

Electrical-field interaction has the capability of introducing loudness distortions and reducing the number of independent frequency channels available for spectral coding (Shannon, 1993). Electrical-field interaction only occurs with simultaneous electrode stimulation (Shannon, 1983, 1985; White et al., 1984; White and Van Compernolle, 1987). Therefore, to avoid electrical-field interactions, speech processing strategies, such as CIS, were developed that stimulate the electrodes with interleaved pulses instead of simultaneous waveforms (Wilson et al, 1991). Although this alleviates the problem of electrical-field interaction, sequential stimulation is not optimal. Transient information can be lost if a set of electrodes corresponding to that frequency region is not stimulated during the time in which the acoustic event occurs. For example, a salient speech feature such as a release burst in a stop consonant could be missed entirely. Therefore, the search for new speech processing strategies continues, and new electrode designs are being developed to take advantage of the benefits of these new strategies.

In the study that follows, speech recognition performance was measured with several new strategies, each varying in the number of simultaneous stimulation channels. As a working hypothesis, we assume that the more simultaneous channels utilized by a speech strategy, the greater the likelihood for electrical-field interactions to occur. The speech strategies were developed in our lab using a research platform known as the Clarion Research Interface (CRI). The effectiveness of each strategy in conveying intelligible speech was measured during acute listening sessions. It was hypothesized that only those patients with the least electrical-field interaction would be able to maintain high levels of speech recognition performance with HAPs and SAS.

The pilot study and the results in the previous chapter demonstrated that the HF+EPS group had the lowest levels of electrical-field interaction and the least interaction spread, respectively. In addition, data from clinics throughout the United States and Europe indicates that patients with the HF+EPS electrode design perform quite well with the simultaneous speech processing strategy, SAS, while those with the ENH electrode design prefer and perform best with sequential strategies, i.e. CIS or PPS (Battmer et al., 2000a, 2000b; Lenarz et al., 1999)[5]. Together, these preliminary results imply that the new Hi-Focus electrode design would be better suited for simultaneous speech processing strategies because patients with this electrode experience reduced levels of electrical-field interaction.

[5] Most of the Hi-Focus patients at the House Ear Institute prefer PPS and most of the Enhanced Bipolar patients prefer CIS

5.2 Method

5.2.1 Subjects

The thirteen subjects from Experiment 1 participated in the speech strategy experiment. However, subject BC’s data was not included in the results because of his extensive duration and degree of hearing loss, which began before the age of 3 (i.e. perilingual hearing loss).

5.2.2 Stimuli and Equipment

The following speech processing strategies were tested on each subject: CIS, PPS, QPS, HAP, and SAS (descriptions of each strategy can be found in Appendix A). The speech strategy software runs on a personal computer and interfaces with an S-Series speech processor through the CRI circuit board. Speech processing strategies are selected and modified with a MATLAB-based user interface developed in our lab. The interface allows the user to enter patient parameters, such as threshold and comfort levels, the electrode coupling mode, and electrode stimulation order.

Speech recognition performance was evaluated for each of these speech strategies with vowel and consonant identification tasks and with an open-set sentence recognition test. The vowels were a subset of those recorded by Hillenbrand et al. (1995). The subset consists of 11 vowels in /hVd/ context: /i/ "heed"; /I/ "hid"; /e/ "hayed"; /ɛ/ “head”; /æ/ “had”; /ʌ/ “hud”; // “hod”; /ɜ/ “herd”; /o/ “hoed”; /ʊ/ “hood”; /u/ “who’d”. The vowel identification task consisted of 3 repetitions of these vowels spoken by 11 adult males, 11 boys, 11 adult females, and 11 girls. Six repetitions of 16 consonants in /aCa/ context, spoken by a single male talker (Shannon et al., 1999), were used for the consonant identification task. Twenty H.I.N.T. sentences (Nilsson et al., 1994) were used for the sentence recognition test. These sentences are spoken by a single male talker, have approximately 6 words, and provide a fair amount of semantic context.

The Speech Identification Utility (SPID), a MATLAB-based software package developed in our lab, was used to administer the speech tests. SPID has a graphical user interface that allows the user to select the speech task and automatically scores the results once the test session is completed. Subjects were seated in a soundproof chamber for the duration of the experiment. All speech material was presented at a 0^{o} azimuth in the soundfield at 65 dB SPL.

5.2.3 Procedure

The clinical programming software developed by Advanced Bionics Corp. (i.e. Software CLINician, or SCLIN) was used to obtain T-levels (thresholds) and M-levels (most comfortable loudness) for each speech processing strategy. These parameters are used to compress the speech signal into the audible dynamic range for each patient and speech strategy. The patient was then fitted with the CRI speech strategy implementation (i.e. CIS, PPS, QPS, HAP, or SAS) and the volume level and sensitivity of the microphone was adjusted to a comfortable listening level. Subjects were familiarized with the test materials prior to the test session.

Testing was divided into five acute listening sessions (20-30 mins) with each speech processing strategy. Each speech strategy session was counterbalanced across subjects to avoid possible order effects. Following the presentation of a vowel or consonant token, the subject was asked to select the button on the computer monitor identifying one of the possible responses. For the sentence recognition task, the subject was asked to repeat as many words in the sentence as possible. The subject was instructed to guess if unsure and no feedback was given during the test session. Results were calculated in percent correct and scored separately for vowel, consonant, and sentence stimuli. The speech tasks were repeated for each speech strategy session, but using a new list of H.I.N.T sentences.

5.3 Results and Discussion

Subject BC's data was excluded from the results described below due to his perilingual hearing loss.

5.3.1. Speech Recognition Performance and Electrode Design

Consonants: Figure 5.3.1.1 illustrates the variability in consonant speech recognition scores across the five speech processing strategies [$F(4,24)=6.41$, $p<.01$]. Consonant speech recognition scores were higher with QPS, CIS, and PPS (QPS: 61.48; CIS: 60.62; PPS: 59.04) and lowest with the HAPs and SAS strategies (HAPs: 45.78; SAS: 40.86%). However, there was no main effect of electrode design nor was there a significant interaction between the speech processing strategy and electrode design.

Vowels: The results for vowel speech recognition are shown in Figure 5.3.1.2. Again there was a main effect of strategy [$F(4,24)=4.56$, $p<.01$] and no effect of the electrode design. The CIS and QPS strategies yielded the highest vowel speech recognition scores (CIS: 53.88%; QPS: 51.99%), PPS and HAPs produced midrange scores (PPS: 49.34%; HAPs: 43.56%), and SAS produced the lowest scores (SAS: 38.07%). A significant interaction was found between the speech strategies and electrode design [$F(8,24)=2.66$, $p<.05$]. The vowel recognition scores of the ENH and ENH+EPS subjects typically decreased as the number of simultaneous channels increased (e.g. SAS & HAPs => QPS, PPS, CIS). In contrast, vowel recognition scores were relatively the same across strategies for the HF+EPS subjects. This pattern was also apparent for the consonant stimuli (Figure 5.3.1.1), although the result did not reach statistical significance.

Sentences: Sentence recognition scores are shown in Figure 5.3.1.3. There were noticeable differences across the speech processing strategies [$F(4,24)=19.28$, $p<.001$]. The PPS, CIS and QPS strategies resulted in the highest speech recognition scores (PPS: 76.87; CIS: 71.26; QPS: 70.57), and SAS and HAPs produced the lowest scores (SAS: 42.47; HAPs: 39.19). As with consonants and vowels, there was no effect of the electrode design. However, the interaction approached significance [$F(8,24)=2.39$, $p=.054$]. The ENH+EPS group had a drastic drop in performance from the sequential speech strategies (i.e. CIS, PPS, QPS) to the fully simultaneous strategy (i.e. SAS), whereas the HF+EPS group had only a slight performance drop.

Taken together, the results from CIS and PPS users demonstrated that speech recognition performance decreased as the number of simultaneous channels increased. Performance was highest for the CIS, PPS, and QPS speech strategies, which have 0, 2, and 4 simultaneous electrodes, respectively. With more than 4 simultaneous channels, there was a noticeable drop in performance.

Previous research has shown that patients generally perform better when tested with the speech processing strategy they preferred when they first received their implant (cite). Therefore, the pattern of results could be explained, at least in part, as a learning effect since performance tended to decrease as a function of stimulation novelty. This is discussed in more detail in Section 5.3.4.

The results for the CIS/PPS users are consistent with the speech processing studies of Wilson et al. (1991) and Boex et al. (1996), in which it was demonstrated that patients achieved significantly higher speech recognition scores with the sequential CIS strategy than with the simultaneous CA strategy. These results were apparent for subjects classified as having high levels of clinical performance as well as for subjects with low levels of performance (Wilson et al., 1995).

If temporal information were the only factor in speech recognition with cochlear implants then SAS would consistently produce higher speech recognition scores than CIS. This is because the amount of temporal information transmitted with a speech strategy is actually greater with more simultaneous channels, e.g. SAS has the capability of providing more temporal information than CIS. However, speech recognition *did not* simply improve as the amount of temporal information increased, but reached a plateau with 4 simultaneous channels (i.e. QPS) and then decreased with a greater number of simultaneous channels. This result indicates that there is a trade-off between improving the temporal resolution with an increasing number of simultaneous channels and introducing distortions from electrical-field interactions. Hence, the QPS strategy defines the upper boundary of “simultaneity” for most of the regular CIS/PPS users in this study.

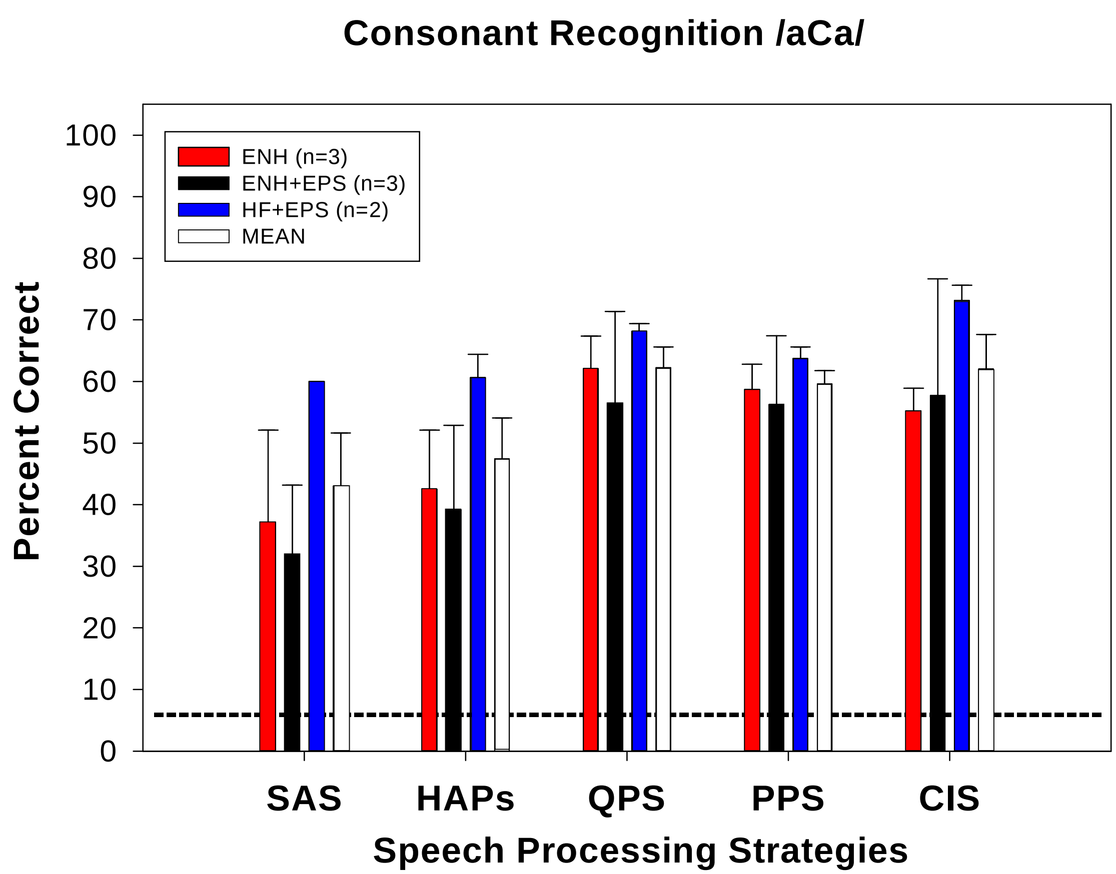


Figure 5.3.1.1: Consonant speech recognition for each of the five speech processing strategies and the three electrode designs. Chance performance was 6% (dashed line).

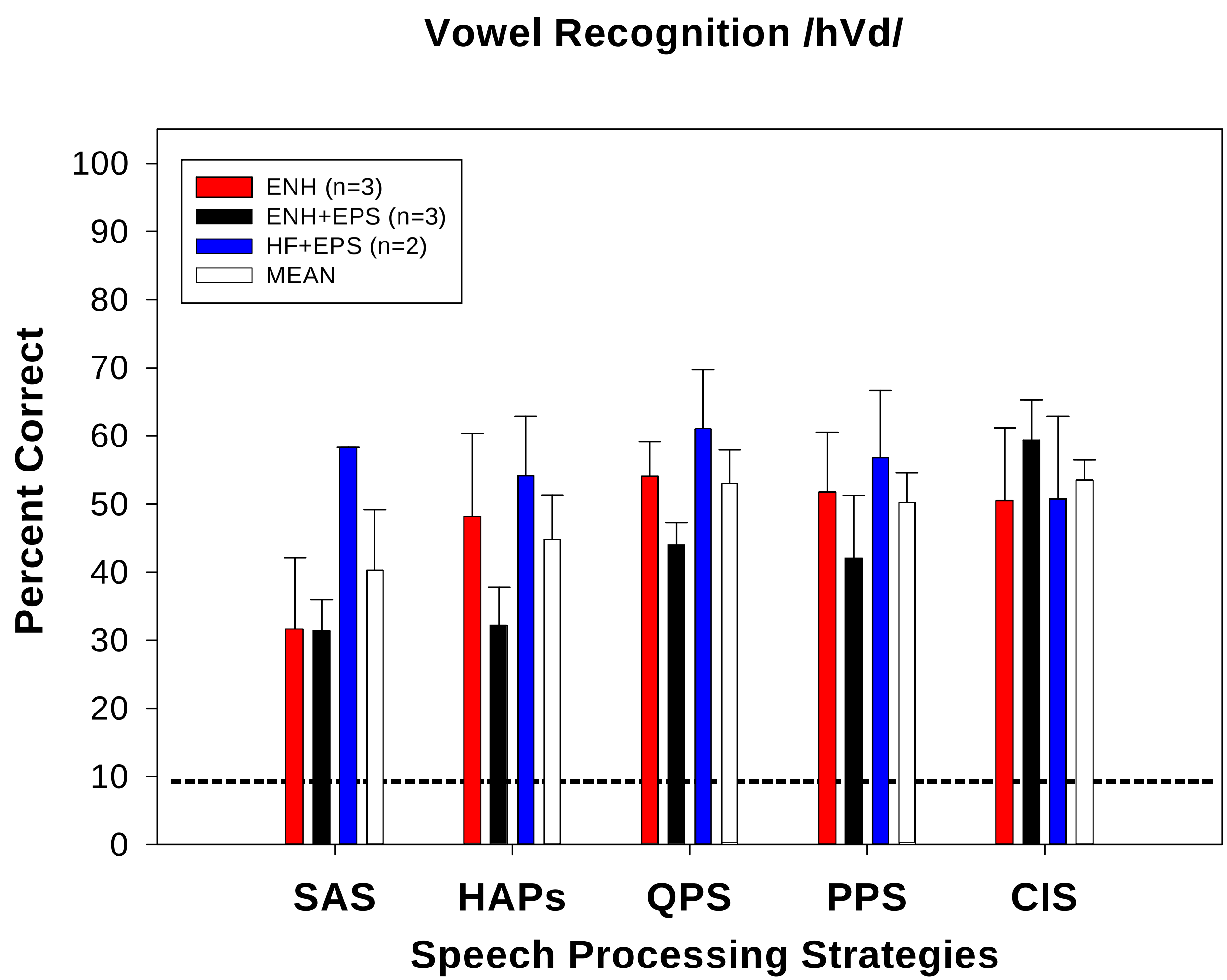


Figure 5.3.1.2 Vowel speech recognition performance for each speech strategy and electrode design. Chance performance was 9% (dashed line).

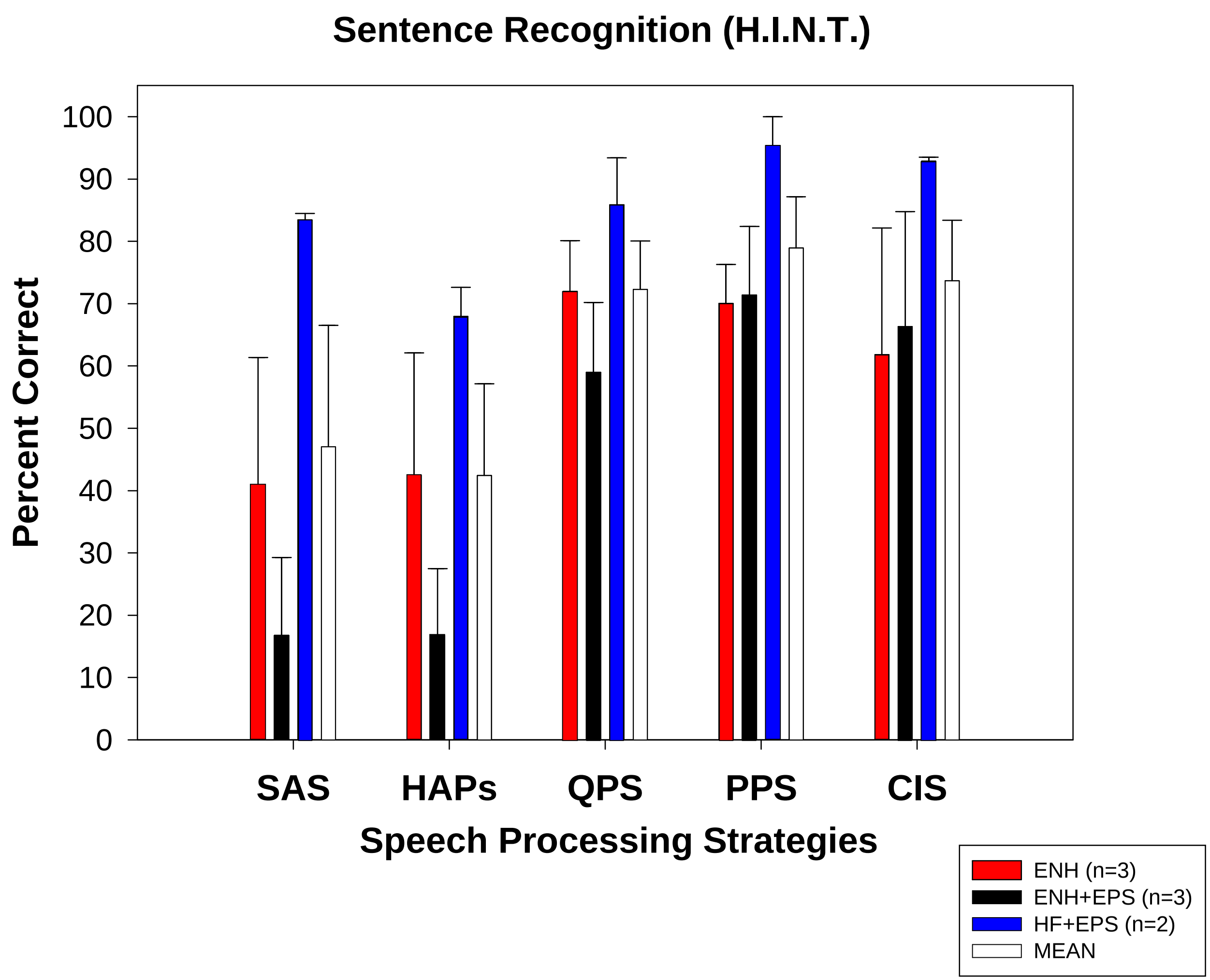


Figure 5.3.1.3: Sentence recognition for each of the speech processing strategies and electrode types.

5.3.2 Speech Processing Strategies and Acoustic Features

The most difficult vowels for the CIS/PPS users to identify were those in "hod" (i.e. /ɑ/) and "head" (i.e. /ɛ/). These results closely mirror those from the study by Loizou, Dorman, and Powell (1998). In both studies, /ɑ/ was most often confused with the vowels in "had" (i.e. /æ/) and "hud" (i.e. /ʌ/). The vowel space proximity between /ɑ/ and /ʌ/ could explain why these two vowels were interchanged. Both vowels have similar F1 and F2 values, although /ɑ/ tends to have a slightly higher F1. The confusion between /ɑ/ and /æ/ is more difficult to explain in terms of the vowel space. Instead, Loizou et al. (1998) suggest that the similar durations of /æ/ and /ɑ/ might have contributed to the substitution. Also, the vowels /ɑ/ and /æ/ have the highest F1 values. These two vowels are mainly differentiated by F2 frequencies. In the /hVd/ context, the F1 transition is relatively static compared to F2 as the frequencies from the vowel move into the final /d/. The confusions between /ɑ/ and /æ/ might have occurred because the F2 transitions into the final consonant were not discernable.

The vowels most affected by switching from the CIS to the SAS strategy were those in "hod" and "heard" (i.e. /ɜ/), while the smallest change was with the vowels in "heed" (i.e. /i/) and "who'd" (i.e. /u/). When the CIS/PPS users were switched to the SAS strategy, they confused /ɜ/ with several vowels having similar F1-F2 values, e.g. sometimes labeling it as /ɛ/ and other times as the vowel in "hoed" (i.e. /o/) or "hood" (i.e. /ʊ/). For the vowel /ɑ/, even more confusions were made with /æ/with the SAS strategy than when the CIS/PPS users were tested with CIS.

For consonant identification, the CIS/PPS users were most successful at identifying the voiceless fricatives /s ʃ/. These two phonemes have energy spread across the frequencies (or electrodes) and, therefore, they have a lot of redundancy. The phonemes /s/ and /ʃ/ are fairly static in that dynamic cues do not contribute much to their identification. They also do not differ substantially in duration. Therefore, the confusion between /s/ and /ʃ/ with the SAS strategy may be due to a loss of spectral information. The most change between CIS and SAS was with the consonants /k/ and /s/. The smallest change was with /g/ and /l/.

5.3.3. Speech Recognition Performance for CIS and SAS Users

Section 3.3.3 demonstrated that the SAS users had lower levels of electrical field interaction than most of the CIS/PPS users. Hypothetically, if electrical field interactions were the primary factor in speech recognition performance, then the SAS users should have approximately the same level of speech recognition performance for all the speech strategies. However, speech recognition performance for the SAS users decreased for speech strategies with fewer simultaneous channels (i.e. PPS and CIS). This was the opposite pattern from the CIS users (see Figures 5.3.1.1 – 5.3.1.3). As mentioned in Section 5.3.1, the rate of stimulation is related to the number of simultaneously stimulated electrodes. A possible explanation for the drop in speech recognition performance for the SAS users is that they had become accustomed to the higher rate of stimulation provided by a fully simultaneous strategy. When the temporal resolution was reduced, so were their speech recognition scores. For example, when SAS users were tested with CIS, /d/ began to sound like /b/. In “ADA”, F2 increases from the initial vowel into the consonant, and decreases from the consonant into the final vowel. In “ABA” the opposite pattern occurs. “ADA” shows more change than “ABA”. Perhaps the direction and frequency extent of the transition is harder to detect with the CIS strategy, which would involve temporal processing. Therefore, learning appears to be a key factor influencing the speech recognition results of both the SAS and CIS users.

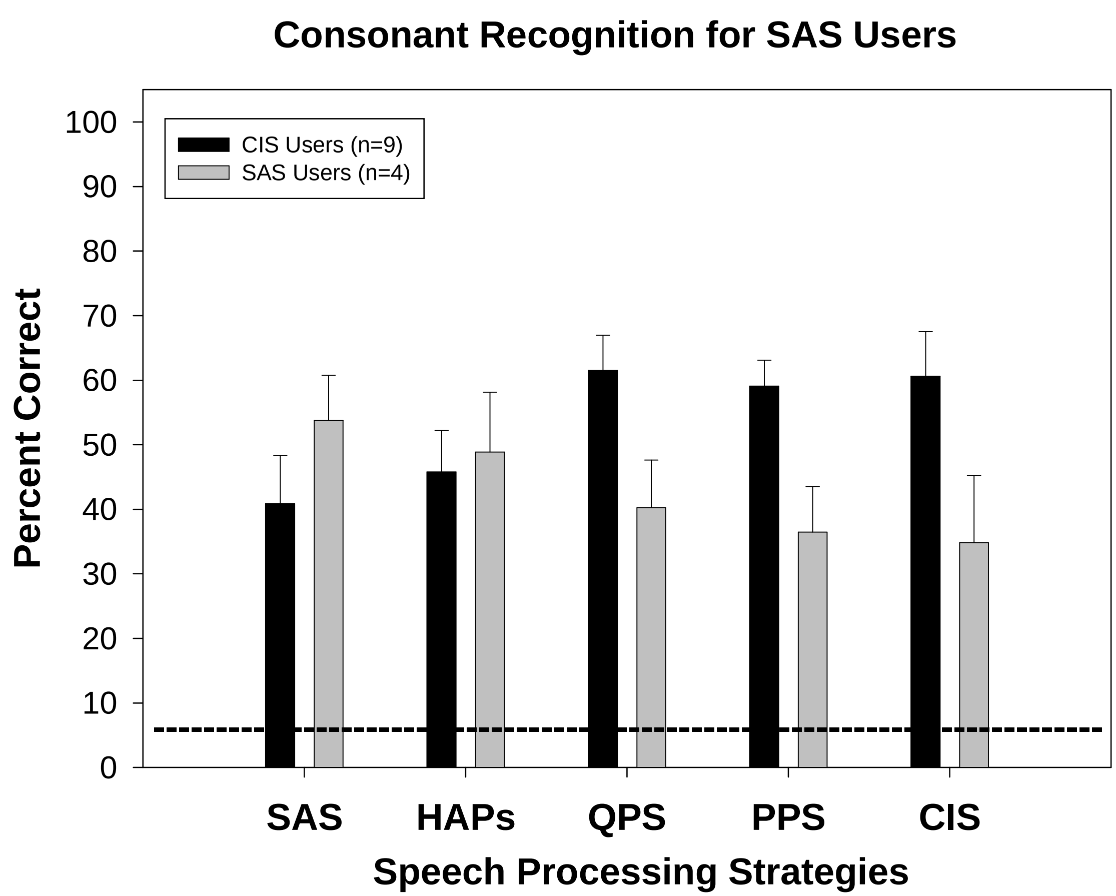


Figure 5.3.3.1: Consonant speech recognition scores are shown for the nine CIS/PPS and 4 SAS users.

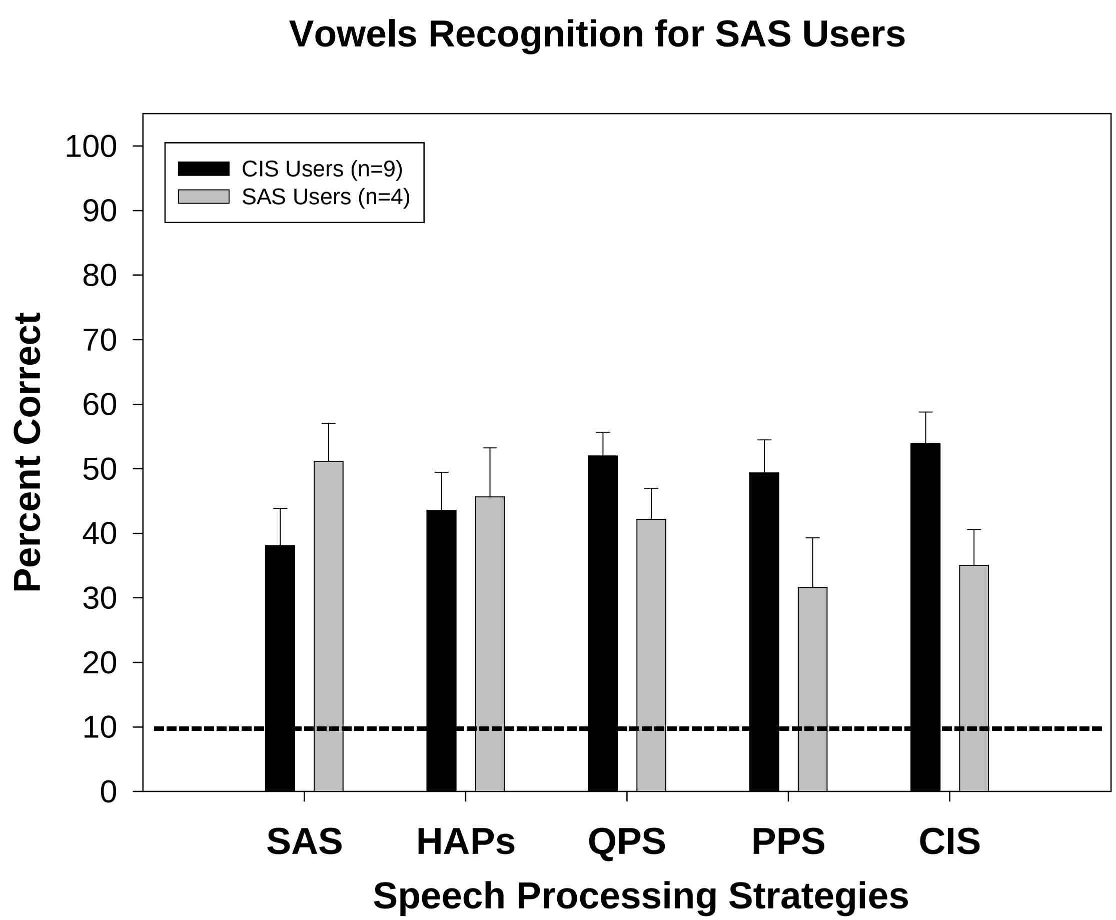


Figure 5.3.3.2: Vowel recognition for SAS and CIS/PPS users.

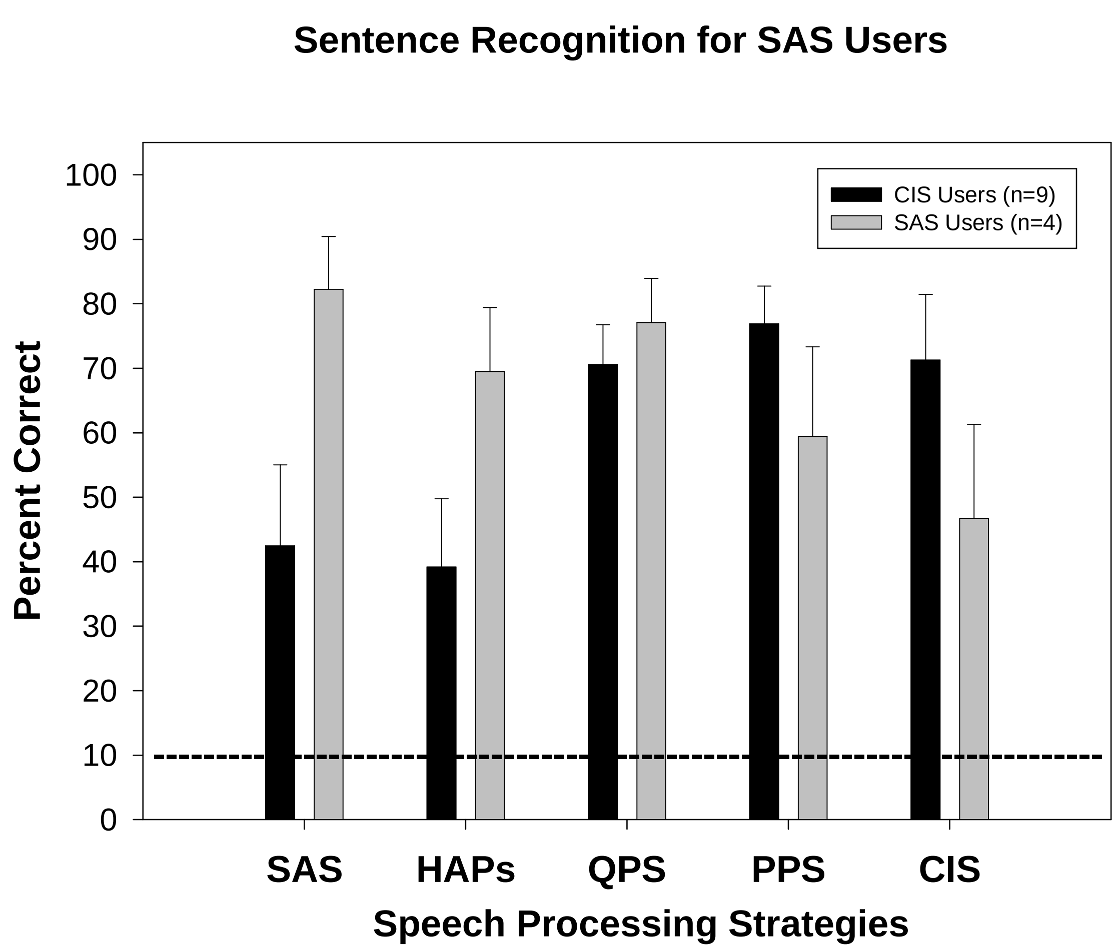


Figure 5.3.3.3: Sentence recognition for CIS/PPS and SAS users.

5.3.4. The Effects of Learning

Novel speech strategies are at a considerable disadvantage with acute listening trials [cite the ACE, SPEAK, CIS studies]. The subjects in this study were generally given only 20-30 minutes of exposure to each of the speech processing strategies, yet had months to years of experience with their own strategy. The performance drop from sequential to simultaneous strategies for CIS users may therefore be due, in part, to the novelty of the stimulus in addition to electrical-field interaction. In contrast, the drop in performance for SAS users from simultaneous to sequential strategies is independent of electrical-field interaction, since electrical-field interactions are less likely to occur with fewer simultaneous channels. For this reason, speech strategy performance for the SAS users was heavily dependent on the novelty of the stimulus and the richness of the information delivered by each strategy (e.g. temporal resolution).

Novel speech strategies exhibiting even a slight improvement from the patient's regular speech strategy have overcome a substantial barrier and may prove to be significantly better with more listening experience. Speech recognition performance was often higher with QPS than with the patient's own strategy. This is of particular interest since the QPS speech strategy is quite different from the strategy used by most of the subjects in this study. This indicates that QPS may be a better speech processing strategy for most of the patients tested in this study.

5.3.5. Speech Processing Strategies: Subjective Descriptions

Prior to testing, the CIS/PPS subjects were asked to describe the sound quality of each speech strategy. Multiple labels were allowed for a particular strategy. Their responses were recorded and compiled in a table as shown below:

	CIS	PPS	QPS	HAPs	SAS	**MEANS**
Vibration	50.00	5.56	12.50	11.11	0.00	**15.83**
Soft	33.33	11.11	0.00	55.56	55.56	**31.11**
Muffled	61.11	33.33	25.00	100.00	83.33	**60.56**
Low Pitch	11.11	11.11	43.75	55.56	33.33	**30.97**
High Pitch	44.44	27.78	6.25	16.67	44.44	**27.92**
Machine-like	38.89	22.22	50.00	72.22	61.11	**48.89**
Static	88.89	33.33	18.75	22.22	0.00	**32.64**
Bubbly	16.67	11.11	0.00	22.22	33.33	**16.67**
Raspy/Gravely	16.67	16.67	25.00	22.22	0.00	**16.11**

Table 5.3.4.1: Percentage of CIS/PPS users who used this label to describe the speech strategy. Numbers in red indicate the most common response for each strategy.

It is important to point out that the CIS/PPS users labeled the HAPs and SAS strategies as “muffled” and “soft”. A plausible reason for this is that the bipolar stimulation used in HAPs and SAS might have made it difficult for the CIS/PPS users to attain sufficient loudness and clarity. This may be why these patients originally preferred the CIS or PPS strategies.

The static heard with the CIS and PPS strategies was described by many of the patients as “bacon frying” or buzzy. This background noise is due to the continuous electrode stimulation at electrical threshold (or T-levels) in the absence of acoustic input (Boyd and Brill, 2000; Pauka, 1989). The static typically becomes imperceptible as the level of stimulation increases. For QPS, SAS, and most of the electrodes in HAPs, the T-levels are set to zero and therefore no background noise is perceived.[6] Typically during a programming session the audiologist would decrease the T-levels for a CIS or PPS strategy if the static persisted.

5.3.6. Future Directions

Soon it will be possible to download experimental speech processing strategies to the patient's speech processor. Once this happens, it will be important to evaluate speech recognition performance over longer periods of time. This will provide valuable information for clinicians and researchers on the time course for adaptation to new speech processing strategies. It is possible that performance with QPS could be improved beyond that found with the patient's own strategy with sufficient listening exposure. Also, each speech processing strategy will have a different rate of speech recognition improvement. Learning rates can provide important clues to identify the critical parameters for speech intelligibility.

---

[6] The T-levels for the CIS or PPS strategy were taken from the patients own MAP. However, PPS T-levels for CIS users and, likewise, CIS T-levels for PPS users were measured with the clinical programming software, SCLIN. T-levels were then decreased 5 clinical units for each electrode.

CHAPTER 6

THE RELATIONSHIP BETWEEN CHANNEL INTERACTION AND SPEECH RECOGNITION PERFORMANCE

6.1 Introduction

In this Chapter, the relationship between electrical-field interaction spread and speech recognition performance is determined with correlation and regression analyses. The results from the pilot study demonstrated that there was a moderate to strong relationship between speech recognition performance and the local electrical interactions arising from simultaneous, *adjacent* electrode stimulation. As electrical-field interaction decreased, speech recognition performance usually increased. This relationship indicates that patients with more electrical-field interaction tend to have more difficulty perceiving speech than patients with less electrical-field interaction. The present study further examines this relationship by correlating speech recognition performance with the *extent* of electrical-field interaction along the cochlea.

A second aim of this Chapter was to investigate the hypothesis that electrical-field interaction is more strongly related to speech recognition using a fully simultaneous speech processing strategy (i.e. SAS) than with a sequential strategy (i.e. CIS). Since sequential stimulation avoids the problem of electrical field summation commonly encountered with cochlear implants, correlations would be relatively weaker for CIS than SAS. Therefore, separate correlations were calculated for speech recognition with CIS and with SAS.

Separate correlations were also made for vowel, consonant, and sentence speech recognition. This produced six separate correlations: vowel recognition with SAS and CIS, consonant recognition with SAS and CIS, and sentence recognition with SAS and CIS. The separate correlations for vowels and consonants were used to determine if interaction spread was more disruptive to the amplitude-envelope or spectral features of speech as measured by the identification of consonants and vowels respectively.

The final aim of the Chapter was to use regression analyses to explore the various factors contributing to speech recognition performance in addition to electrical-field interaction. These factors are as follows: the duration of hearing loss, defined as the amount of time the patient experienced a mild to severe hearing loss; duration of deafness, defined as the amount of time with a profound hearing loss prior to obtaining a cochlear implant; duration of cochlear implant use; and the patient's age.

6.2 Results and Discussion

The following sections discuss the results from the CIS/PPS users only, since there was an insufficient number of SAS users (n=4) for the correlation or regression analyses.

6.2.1. The Relationship between Electrical-field interaction and Speech Recognition Performance

Separate correlations were calculated for CIS and SAS vowel, consonant, and sentence recognition. Since CIS uses monopolar stimulation and SAS uses bipolar stimulation, CIS speech recognition performance was correlated with *monopolar* interaction spread (as measured by the area under the interaction spread curve) and SAS speech recognition performance was correlated with *bipolar* interaction spread.

Figure 6.2.1 shows the results for each condition. A moderate to strong negative correlation was found between interaction spread and SAS speech recognition, whereas a relatively weak relationship was found between interaction spread and CIS speech recognition. The SAS correlations are consistent with the hypothesis that simultaneous stimulation will only yield high speech recognition scores when the amount of electrical-field interaction is absent or substantially reduced.

There are two reasons why the SAS conditions produced higher correlations than the CIS conditions. First, CIS is a sequential strategy, and because each electrode is stimulated in turn, CIS avoids electrical field summation. It is quite peculiar then that the CIS correlations from the pilot study were so high! Possible reasons for this discrepancy are discussed in Section 6.2.4. The second reason for the higher SAS correlations is that simultaneous monopolar stimulation produces roughly the same amount of electrical-field interaction, yet there was a wide range of speech recognition abilities across subjects.

Relationship Between Interaction Spread and Speech Recognition

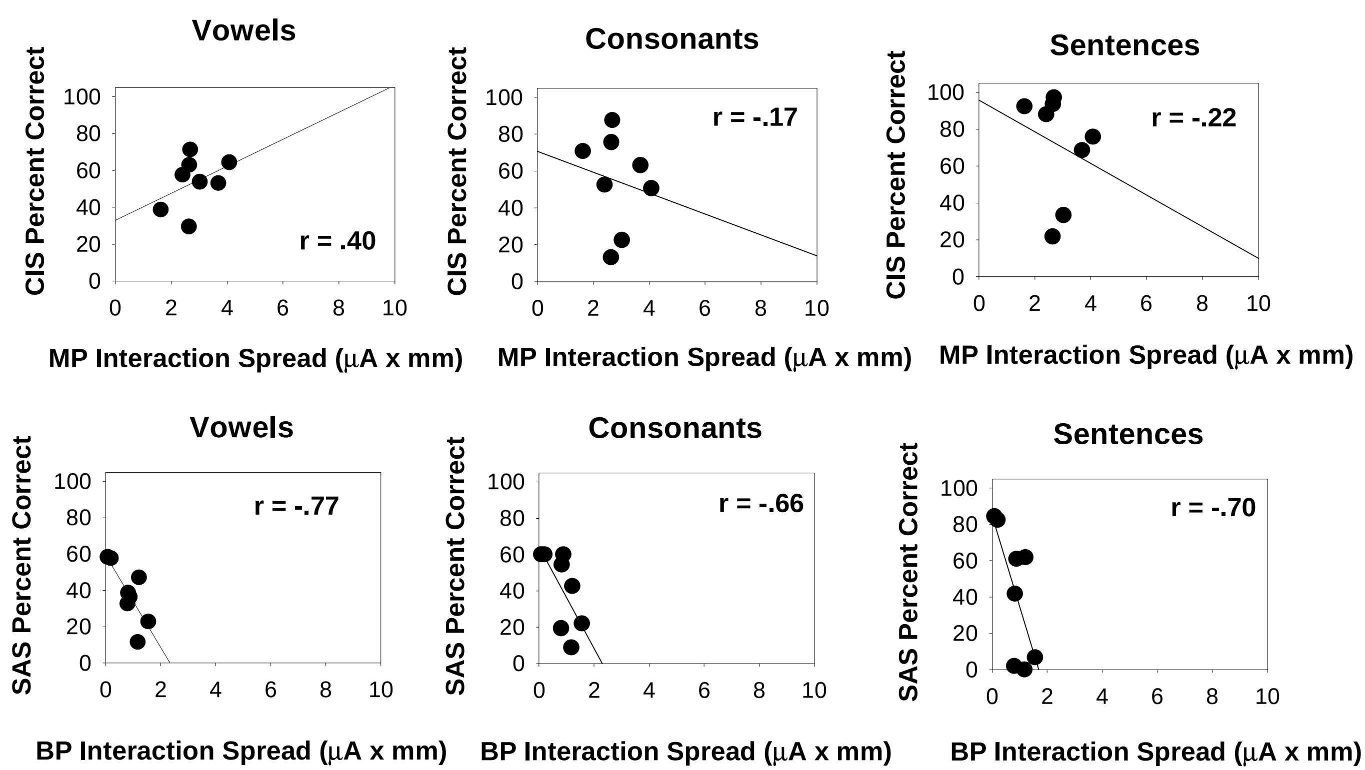


Figure 6.2.1: Correlations show the relationship between CIS and SAS speech recognition scores and interaction spread. Interaction spread, shown on the x-axis, is the area under the interaction spread curve for each subject. The y-axis is the percent correct score for the CIS (upper panels) or SAS strategy (lower panels). Monopolar interaction spread was used for CIS correlations and bipolar interaction spread was used for SAS correlations. Vowel correlations are shown in the left panel, consonant correlations are in the middle panel, and sentence correlations are in the right panel.

6.2.2. The Relationship Between Acoustic Features and Channel Interaction

Slightly stronger correlations were found for vowels than consonants. In addition, Figure 6.2.2 demonstrates that consonant place-of-articulation (e.g. labial, alveolar, velar) was more strongly correlated with interaction spread than the consonant features of voicing (e.g. voiced and unvoiced) and manner (e.g. stop, fricative, affricate, nasal or glide). These results indicate that electrical-field interaction is more harmful to speech stimuli that rely heavily on the processing of spectral cues. Electrical-field interaction may smear the boundaries between spectral peaks or introduce aberrant peaks, making it difficult to distinguish between neighboring formants. The consonant features of voicing and manner were relatively well resolved in the presence of electrical-field interaction since fairly course spectral cues and temporal envelope information are sufficient for identifying these features (Fishman et al., 1997; Rabinowitz and Eddington, 1995; Shannon et al., 1995; Van Tasell et al., 1987, 1992).

Voicing showed a stronger relationship with electrical-field interaction than manner. Most of the voicing errors occurred with fricatives. When the CIS/PPS users were tested with CIS, voicing errors occurred within the same place category (e.g. velar /g/ confused with /k/; labio-dental /v/ confused with /f/). When tested with SAS, voicing errors were again confused within the same place category for alveolar /s/ with /z/, however, the confusions were even more extreme where manner, place, and voicing cues were disrupted (e.g. bilabial stop /b/ with labio-dental fricative /v/; postalveolar affricate /dʒ/ with alveolar stop /t/). The acoustic cues for consonant voicing depend on F1 information and the relative timing of events across frequencies. If electrical-field interaction blurs the boundaries between formants, then it would be difficult for a listener to detect the onset or offset of a neighboring formant.

Relationship Between Electrical-field interaction and Consonant Features

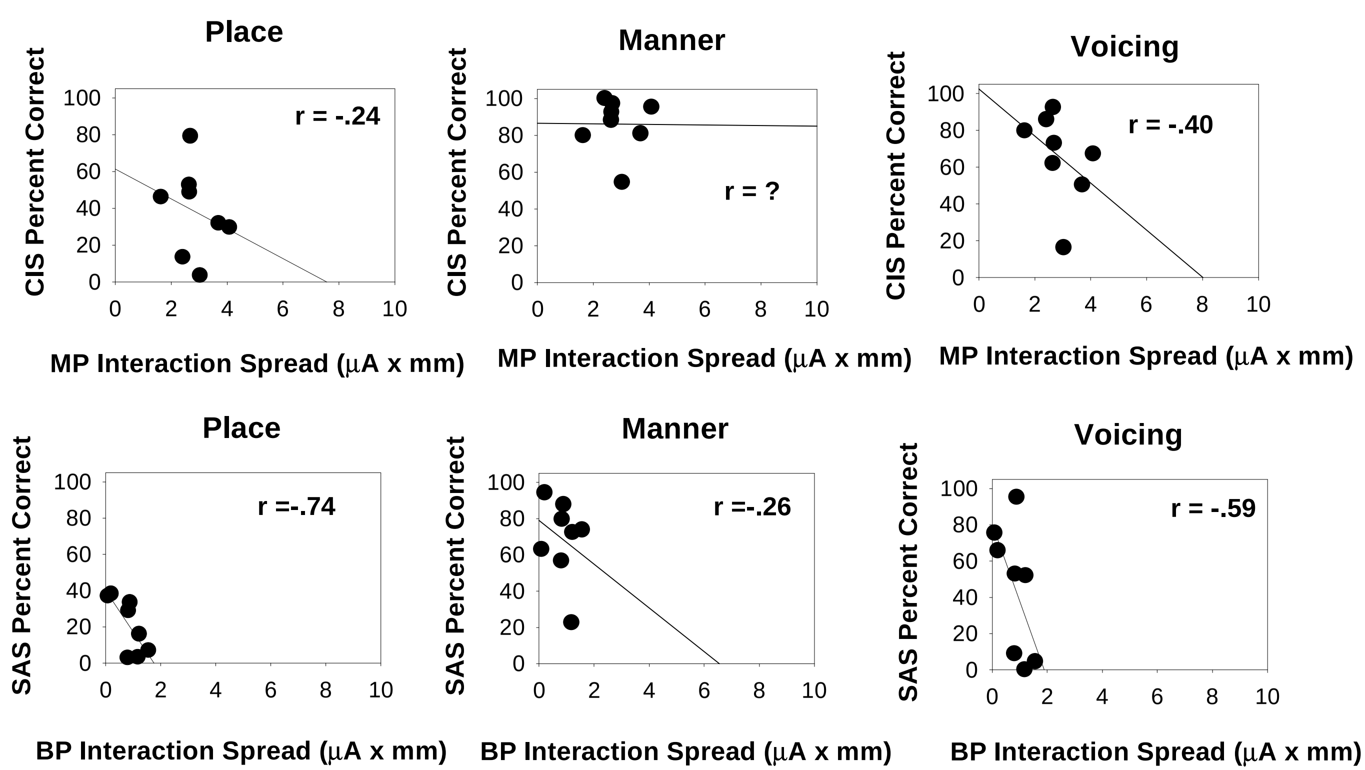


Figure 6.2.2. Separate correlations are shown for the consonant features of place (left panel), manner (middle panel), and voicing (right panel). Correlations between monopolar interaction spread and the three CIS consonant features are shown in the top panels, while correlations between bipolar interaction spread and SAS consonant features are shown in the bottom panels.

6.2.3. Other Factors Contributing to Speech Recognition Performance

Multiple regression analyses examined the relative contributions of electrical-field interaction, age, duration of deafness, duration of hearing loss, and cochlear implant experience to CIS and SAS speech recognition performance. For the SAS conditions, electrical-field interaction had the greatest contribution to speech recognition performance, followed by the duration of deafness. The factors combined accounted for 47% of the variance in SAS speech recognition scores for consonants, 79% for vowels, and 55% for sentences. For the CIS conditions, the factor(s) having the greatest influence on speech recognition performance varied depending on the speech task. Electrical-field interaction and the duration of hearing loss had the greatest influence on CIS vowel recognition scores, and the factors combined accounted for 92% of the variance. Almost all the factors had an equal contribution to CIS sentence scores, accounting for 33% of the variance. In contrast, only 17% of the variance in CIS consonant scores could be explained.

6.2.4. Comparisons with Pilot Data

The correlations between electrical-field interaction and speech recognition performance were slightly higher with the pilot data than in the present study. Two reasons for this result are discussed below.

First of all, the Hi-Focus subjects were not included in the correlations for the pilot study. When the Hi-Focus data was included, the correlations dropped. This may be related to the finding that the Hi-Focus patients typically did not have electrical-field interactions even with adjacent masker+probe pairs.

Second, CIS is sequential, but has 4μsec of overlap between adjacent pulses. Therefore, CIS would only be subjected to electrical-field interactions between adjacent electrodes. In the pilot data, the speech scores were correlated with the magnitude of channel interaction between ADJACENT electrodes. In the present study, we're correlating speech scores with interaction spread, or the pattern of interaction as a function of masker, probe separation. Therefore, the magnitude of channel interaction occurring locally would be a better predictor of CIS speech recognition than the area of interaction spread.

CHAPTER 7

General Discussion

When cochlear implants were first introduced commercially in the mid 1960s most clinicians and researchers hoped that the prosthesis would help a limited number of patients achieve some degree of hearing sufficient for detecting speech and environmental sounds. Since that time, the combined efforts from many disciplines, including bioengineering, speech and hearing science, and otolaryngology, have led to developments with cochlear implants that have far exceeded these expectations. Today the cochlear implant has become the most successful implanted prosthesis to interface with the human sensory system.

Although the number of “star” performers has increased since implants were first introduced, there still remains substantial variability across patients. The variability in patient performance can be attributed to several factors, many of which were examined in this dissertation, including the duration of deafness, cochlear implant experience, electrical-field interactions, speech processing strategies, and electrode design. Electrode interaction was clearly associated with speech recognition performance and therefore needs to be taken into consideration during the evaluation and fitting of cochlear implants to bring about the most benefit for each individual patient.

The type of stimulation used in speech processing strategies can also influence sppech intelligibility. Speech processing strategies that stimulate multiple electrodes simultaneously (i.e. CA and SAS) are not effectively used by a substantial number of patients. The results from this dissertation indicate that electrical-field interactions appear to be a major limiting factor. As such, the range of speech processing strategies available to each patient is restricted and the potential benefits of simultaneous strategies (or hybrids such as HAPs) will not be recognized until electrical-field interactions can be further reduced.

The dissertation also demonstrated that electrode design modifications, such as the silastic buffers and electrode positioner of the Hi-Focus electrode, reduce electrical-field interactions considerably. Recent studies examining electrodes that achieve modiolar proximity and their effect on current levels have motivated other cochlear implant manufactures (i.e. MedEl and Cochlear Corp.) to join in the search for new electrode designs and positioning techniques to reduce current spread and electrical-field interactions.

The coupling mode used in simultaneous stimulation is also a major contributor to electrical-field interactions. This study demonstrated that more focused stimulation, i.e. bipolar stimulation, is required with simultaneous speech strategies to avoid interactions. Researchers are currently investigating ways to produce even more highly focused current fields with new coupling modes, such as the quadrupolar mode, in hopes that they may further reduce electrical-field interactions (Jolly et al., 1996).

The subjects in these experiments clearly showed a preference for a particular speech processing strategy, both in their level of performance and the subjective descriptions for each of the strategies. In fact, subjects showed the best performance with the speech processing strategies most similar to their own. For the CIS/PPS subjects, lower speech recognition performance occurred when the electrode stimulation became more simultaneous, e.g. the highest speech recognition scores were obtained with the CIS strategies, midrange scores occurred for the HAP strategy, and the lowest scores were produced with the SAS strategy. In contrast, the SAS users showed the opposite pattern - higher speech recognition scores with the SAS strategy and the lowest scores with the CIS strategy. These results can be partly attributed to electrical-field interaction (for the PPS/CIS users) and to stimulation rate (for the SAS users). However, it is presently unclear whether the pattern of performance was strictly the result of the electrical stimulation or also influenced by a learning effect, since the subjects had more listening experience with their own strategy. Longer listening trials with each of the speech strategies may help answer this question.

It was claimed that if a patient shows higher performance with a novel speech strategy, then it can be inferred that the new strategy is better suited for that patient. The QPS strategy was one strategy that appeared to hold promise. Speech recognition scores were either close to those with the patient's own strategy or better. Therefore, QPS is one strategy that should be evaluated with prolonged exposure to determine if the higher rate of stimulation is beneficial.

This dissertation has combined psychophysical, signal processing, and speech recognition tasks to demonstrate the relationship between electrical-field interactions and speech recognition performance. Although previous studies in this area have addressed electrical-field interactions in cochlear implants, none have combined these techniques to measure the critical link between speech intelligibility and electrode interactions. These experiments were the first step towards uncovering the detrimental effects of electrical-field interactions on speech recognition performance. Future studies will be needed to directly address how to take electrical-field interactions into account when fitting a cochlear implant.

Appendix A

Continuous Interleaved Sampler (CIS)

CIS is a nonsimultaneous, pulsatile speech processing strategy that uses a monopolar coupling mode. The CLARION CIS strategy delivers pulsatile stimulation at the rate of 6,500 pulses per second (pps) with a pulse width of 75 μs/phase. The CIS strategy stimulates the electrodes sequentially from apex-to-base (i.e. 1,2,3,4,5,6,7,8). However, with this type of stimulation, temporal interactions may occur as the result of current spread. For example, if electrode 4 is stimulated, it's current might spread to and effect the neural response to electrodes 3 and 2 in the apical direction and electrodes 5 and 6 in the basal direction. When electrode 3 is subsequently stimulated, the underlying neural population may be in recovery from the previous stimulation of electrode 4 and may only respond in a weak manner or not at all. With the CIS strategy, the electrodes can also be stimulated in a non-sequential manner, thereby reducing the potential for both temporal and spatial interactions. Staggered stimulation may be the optimal choice for CIS users if electrical-field interaction is only a concern for neighboring electrodes. An example of staggered stimulation would be as follows: 1,5,2,7,3,8.

Paired Pulsate Sampler (PPS)

The PPS strategy is a partially simultaneous, pulsatile speech strategy that shares some features with CIS (e.g. monopolar stimulation). However, instead of stimulating one electrode at a time, two electrodes are stimulated simultaneously. Three electrodes separate simultaneous electrode pairs in the default PPS design. For example, electrodes 1 and 5 are stimulated together, then 2 and 6, 3 and 7, and so on. Because of the simultaneous stimulation, the rate is increased to about twice that of the CIS strategy (PPS: 13,300 pps; CIS: 6500 pps). The same pulse width is used for both strategies (75 μs/phase). The increase in the rate of stimulation provides better temporal resolution, which may improve speech recognition of transient speech events (Loizou, 1999).

Quadruple Pulsatile Strategy (QPS)

The QPS strategy is also a partially simultaneous, pulsatile strategy like PPS, but stimulates 4 electrodes at a time instead of two. One electrode separates simultaneous channels in the QPS strategy. All odd numbered electrodes are stimulated together, i.e. electrodes 1,3,5, and 7, then all the even numbered electrodes are stimulated together, i.e. 2,4,6, and 8. The QPS strategy has the same pulse width as PPS and CIS (i.e. 75 μs/phase) and stimulates at a pulse rate of 3.3 Kpps. QPS has a higher stimulation rate than either PPS or CIS. QPS uses monopolar stimulation.

Simultaneous Analog Stimulation (SAS)

The SAS speech processing strategy delivers simultaneous analog waveforms to channels 1-7. The rate of stimulation is 91,000 pps. SAS uses the bipolar stimulation mode. For the standard Clarion electrode an "enhanced bipolar" coupling mode is used, with an active medial electrode referenced to the lateral electrode of the adjacent, basal electrode pair. The Hi-Focus electrode uses two adjacent medial electrodes to generate a bipolar current field. Therefore, electrode 8 is excluded from SAS stimulation since there is no basal electrode beyond electrode 8 for coupling.

Hybrid Analog Pulsatile (HAP)

The HAP strategy combines CIS and SAS speech processing, using pulsatile waveforms for CIS electrodes and analog waveforms for SAS electrodes. For this study, the HAP strategy delivers SAS stimulation on apical channels 1-5 and CIS stimulation to basal channels 6 and 7. The potential for electrical-field interactions, due to poor nerve survival in the basal end of the cochlea, can be avoided with sequential stimulation in those channels. The advantage of using a simultaneous, fast rate stimuli on apical channels is to preserve temporal and spectral information in the low and midrange frequencies. The rate of the HAP strategy is related the number of simultaneous channels using SAS.

References

Aitkin, L (1990). *The auditory cortex*, (Chapman and Hall, London).

Abbas, PJ (1993). "Electrophysiology," in *Cochlear Implants: Audiological foundations*, edited by RS Tyler, (Singular Publishing, San Diego), pp. 317-355.

Armstrong-Bednall, G, Goodrum-Clarke, K, Stollwerck, L, Nunn, T, Wei, J, Boyle, P (1999). "Clarion paired pulsatile sampler (PPS): User preference and performance," Presented at the 1999 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.

Battmer, RD, Gnadeberg D, von Wallenberg E, (1993). "A study of monopolar and bipolar stimulation modes with a modified Nucleus mini-22 cochlear implant," Adv Otorhinolaryngol 48, 9-16.

Battmer, RD, Zilberman, Y, Haake, P, Lenarz, T (1999). "Simultaneous Analog Stimulation (SAS) – Continuous Interleaved Sampler (CIS) pilot comparison study in Europe," Ann Otol Rhinol Laryngol 108, 69-73.

Battmer, RD, Goldring, JE, Kuzma, J, Lenarz, T (2000a). "New CLARION Hi-Focus Intracochlear Electrode: Long-term clinical results," Presented at The 6th International Cochlear Implant Conference. Miami Beach, FL.

Battmer, RD, Goldring, JE, Kuzma, J, Lenarz, T (2000b). "Intraoperative measures and post-operative clinical results with new modiolus hugging electrodes and simultaneous analog stimulation (SAS) in young children," Presented at The 6th International Cochlear Implant Conference. Miami Beach, FL.

Berliner, KI and Eisenberg, LS (1987). "Our experience with cochlear implants: have we erred in our expectations?," Am J Otol 8, 222-229.

Berliner, KI, Tonokawa, LL, and Dye, LM, et al (1989a). "Open-set speech recognition in children with a single-channel cochlear implant.", Ear Hear 10, 237-242.

Berliner, KI, House, WF, Tonokawa, LL (1989b). "Open-set speech recognition by children with a single-channel cochlear implant,", Trans Am Otol Soc, 50.

Boex, C, de Balthasar, C, and Pelizzone, M (1999). "Electrode interactions in Clarion subjects," Presented at the 1999 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.

Boex, C, Pelizzone, M, and Montandon, P (1996). "Speech recognition with a CIS strategy for the Ineraid multichannel cochlear implant," Am. J. Otol. 17, 61-68.

Boothroyd, A., Mulhearn, B., Gong., J., and Ostroff, J. (1996). "Effects of spectral smearing on phoneme and word recognition," J. Acoust. Soc. Am. 100, 1807-1818.

Boyd, PJ and Brill, S (2000). "Psychoacoustic measures of electrical threshold interactions in the CIS strategy," Presented at The 6th International Cochlear Implant Conference. Miami Beach, FL.

Burns, EM and Viemeister, NF (1976). “Nonspectral pitch,” J. Acoust. Soc. Am. 60, 863-869.

Burns, EM and Viemeister, NF (1981). “Played-again SAM: Further observations on the pitch of amplitude-modulated noise,” J. Acoust. Soc. Am. 70, 1655-1660.

Busby, PA, Whitford, LA, Blamey, PJ, Richardson, LM, and Clark, GM. (1994). “Pitch perception for different modes of stimulation using the cochlear multiple-electrode prosthesis,” J. Acoust. Soc. Am. 95, 2658-2669.

Carhart, R and Jerger, JF (1959). “Preferred method for clinical determination of pure-tone thresholds,” Journal of Speech and Hearing Disorders 24(4), 330-345.

Chatterjee, M and Shannon, RV. (1998). “Within-channel gap detection using dissimilar markers in cochlear implant listeners,” J. Acoust. Soc. Am. 103, 2515-2519.

Chouard, CH and MacLeod, P (1976). “Implication of multiple intracochlear electrodes for rehabilitation of total deafness: Preliminary report,” Laryngoscope 86: 1743-1751.

Cotter, NE (1986). “Modeling of auditory prostheses,” Doctoral dissertation, Stanford University, Stanford, CA. (Stanford Electronics Laboratories Technical Report No. G906-8).

Djourno A and Eyries C (1957). “Prosthese auditive par excitation electrique a distance du nerf sensoriel a l’aide d’un bobinage inclus a demeure,” Presee Med 65, 14-17.

Donaldson, GS and Nelson, DA. (2000). "Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies," J. Acoust. Soc. Am. 107(3), 1645-1658.

Dorman, MF (1993). "Speech perception by adults," in *Cochlear Implants: Audiological foundations*, edited by RS Tyler, (Singular Publishing, San Diego), pp. 145-190.

Dorman, MF, Loizou, P, and Rainey, D. (1997). "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. Am. 102, 2403-2411.

Dorman, MF and Loizou, PC (1998a). "The identification of consonants and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normal-hearing subjects using simulations of processors with two to nine channels," Ear and Hear. 19(2), 162-166.

Dorman, MF, Loizou, P, and Tu, Z. (1998b). "The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels," J. Acoust. Soc. Am. 104(6), 3583-3585.

Eddington, DK, Dobelle, WH, Brackmann, DE, Mladejovsky, MG and Parkin, JL (1978). "Auditory prostheses research with multiple channel intracochlear stimulation in man," Ann Otol Rhinol Laryngol 87(suppl 53), 5-38.

Erber, NP (1979). "Speech perception by profoundly hearing-impaired children," J. Speech Hear. Disord. 44, 255-270.

Fishman, K, Shannon, RV, and Slattery, WH. (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, RV, Wang X (1998). “Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing,” J. Acoust. Soc. Am. 104(6): 3586-3596.

Fu, Q-J and Shannon, RV (1999). “Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing,” J. Acoust. Soc. Am. 105(3): 1889-1900.

Gantz, B, Tyler, R, Knutson, J, Woodworth, G, Abbas, P, McCabe, B, Hinrichs, J, Tye-Murray, N, Lansing, C, Kuk, F, and Brown, C (1988). “Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance,” Laryngoscope 98, 1100-1106.

Greenwood, DD (1990). “A cochlear frequency-position function for several species - 29 years later,” J. Acoust. Soc. Am. 87, 2592-2605.

Hanekom, JJ and Shannon, RV. (1998). “Gap detection as a measure of electrode interaction in cochlear implants,” J. Acoust. Soc. Am. 104(4), 2372-2384.

Hillenbrand, J, Getty, L, Clark, M, and Wheeler, K (1995). “Acoustic characteristics of American English vowels,” J. Acoust. Soc. Am. 97(5), 3099-3111.

House, WF (1994). “Cochlear implants. It’s time to rethink,” Am. J. Otol. 15, 573-587.

Jolly, CN, Spelman, FA, and Clopton, BM (1996). “Quadrupolar stimulation for cochlear prostheses: modeling and experimental data,” IEEE Trans Bio Eng 43(8), 857-865.

Leake, PA, Snyder, RL, Rebscher, SJ, Moore, CM, Vollmer, M (2000). “Plasticity in central representations in the inferior colliculus induced by chronic single- vs. two-channel electrical stimulation by a cochlear implant after neonatal deafness,” Hear Res 147, 221-241.

Lenarz, T, Battmer, R, Kuzma, J, Maltan, A (1999). “The HiFocus electrode system,” Presented at the 1999 Conference on Implantable Auditory Prosthesis, Pacific Grove, CA.

Lenhardt E, Gnadeberg D, Battmer RD, von Wallenberg E (1992). “Experience with the cochlear miniature speech processor in adults and children together with a comparison of unipolar and bipolar modes,” Otorhinolaryngology 54, 308-313.

Lawson, DT, Wilson, BS, Zerbi, M, and Finley, CC (1996). “Speech processors for auditory prostheses,” Third Quarterly Progress Report (NIH Contract N01-DC-5-2103).

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

Lim, HH, Tong, YC, and Clark, GM (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.

Loizou, PC and Poroy, O (1999). “The effect of parametric variations of cochlear implant processors on speech understanding,” J. Acoust. Soc. Am., 108(2), 1-13.

Loizou, PC, Dorman, MF, and Powell, V (1998). “The recognition of vowels produced by men, women, boys, and girls by cochlear implant patients using a six-channel CIS processor,” J. Acoust. Soc. Am., 103(2), 1141-1149.

Merzenich, MM, Schindler, DH, White, MW (1974). “Symposium on cochlear implants. II. Feasibility of multichannel scala tympani stimulation,” Laryngoscope. 84: 1887-1893.

Miller, G. and Nicely, P. (1955). “An analysis of perceptual confusions among some English consonants,” J. Acoust. Soc. Am. 27, 338-352.

Miyamoto, RT, Osberger, MJ, Cunningham, L, Kirk, KI, Myres, WA, Robbins, AM, Kessler, K (1994). “Single-channel to multichannel conversions in pediatric cochlear implant recipients,” Am. J. Otol. 15(1), 40-45.

Moller, AR (1999). “Review of the roles of temporal and place coding of frequency in speech discrimination,” Acta Otolaryngol (Stockh), 119, 424-430.

Nilsson, M, Soli, S, and Sullivan, J (1994). “Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise,” J. Acoust. Soc. Am. 95, 1085-1099.

Osberger, MJ and Fisher, L (1999). “SAS-CIS preference study in postlingually deafened adults implanted with the CLARION® cochlear implant,” Ann Otol Rinol Laryngol 108, 74-79.

Osberger, MJ, Koch, D, Fisher, L, Zimmerman-Phillips, S (1999). “Clinical results in patients implanted with the Clarion Electrode Positioning System,” Presented at the 1999 Conference on Implantable Auditory Prosthesis, Pacific Grove, CA

Pauka, CK (1989). “Place-pitch and vowel-pitch comparisons in cochlear implant patients using the Melbourne-Nucleus cochlear implant,” J. Laryngol. Otol., Suppl. 19, 1-31.

Pfingst, BE (1988). “Comparisons of psychophysical and neurophysiological studies of cochlear implants, Hear. Res. 34, 243-252.

Pfingst, BE and Rush, NL (1987). “Discrimination of simultaneous frequency and level changes in electrical stimuli,” Ann of Otol Rhin Laryng 96(Suppl. 128), 34-37

Rabinowitz, WM and Eddington, DK (1995). “Effects of channel-to-electrode mappings on speech reception with the Ineraid cochlear implant,” Ear and Hearing 16(5), 450-458.

Rebscher, SJ, Heilmann, M, Talbot, N, Bruszewski, W, and Merzenich, M. (1994). “Studies on pediatric auditory prosthesis implants,” Sixteenth Quarterly Progress Report, NIH Contract N01-DC-2401.

Rosen, S (1989). “Temporal information in speech and its relevance for cochlear implants,” in Cochlear Implant: Acquisitions and controversies, edited by B Fraysee and N Cochard, (Toulouse Implant Conference Proceedings, Toulouse), pp. 3-26

Rosen, S and Smith, (1988). “Temporally-based auditory sensations in the profoundly hearing-impaired listener,” in *Basic Issues in Hearing*, edited by H. Duifhuis, JW Horst, and HP Wit, (Academic Press, London), pp. 431-449.

Rosen, S, Faulkner, A, and Wilkinson, L (1999). “Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants,” J. Acoust. Soc. Am. 106(6), 3629-3636.

Rubinstein, JT, Parkinson, WS, Lowder, MW, Gantz, BJ, Nadol Jr., JB, Tyler, RS. (1998). “Single-channel to multichannel conversions in adult cochlear implant subjects,” Am. J. Otol. 19: 461-466.

Sachs, MB and Young, ED (1979). “Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate,” J. Acoust. Soc. Am. 66: 470-470.

Schindler, RA, Kessler, DA, Barker, M (1995). “Clarion patient performance: an update on the clinical trials,” Ann. Otol. Rhinol. Laryngol. 104(suppl. 166), 269-72.

Shannon, RV. (1983a). “Multichannel electrical stimulation of the auditory nerve in man: I. Basic Psychophysics,” Hearing Res. 11, 157-189.

Shannon, RV. (1983b). “Multichannel electrical stimulation of the auditory nerve in man: II. Channel interaction,” Hearing Res. 12, 1-16.

Shannon, RV. (1985).”Loudness summation as a measure of channel interaction in a cochlear prosthesis,” in *Cochlear Implants*, edited by RA Schindler and MM Merzenich, (Raven Press, New York), pp. 323-333.

Shannon, RV (1992). “Temporal modulation transfer functions in patients with cochlear implants,” J. Acoust. Soc. Am. 91, 1974-1982.

Shannon, RV, Zeng, F-G, and Wygonski, J. (1992). “Speech recognition with only temporal cues,” in *Speech Processing: From the cochlea to language*, edited by MEH Schouten, (Mouton-Gruyer, Berlin).

Shannon, RV (1993). “Psychophysics,” in *Cochlear Implants: Audiological foundations*, edited by RS Tyler, (Singular Publishing, San Diego), pp. 357-388.

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

Shannon (1999) (for consonants)

Shepherd, RK, Hatsushika, S, and Clark, GM (1993). “Electrical stimulation of the auditory nerve: The effect of electrode position on neural excitation,” Hear Res 66, 108-120.

Stollwerck L, Goodrum-Clarke K, Lynch C, Armstrong-Bednall G, Nunn T, Markoff L, Mens L, McAnallen C, Wei J, Boyle P, George C, Zilbermann Y (1999). “Speech processing strategy preferences among 55 European CLARION cochlear implant users,” Presented at EFAS, Oulu, Finland.

Stickney, GS, Shannon, RV, Opie, JM, and Assmann, PF. (2000). “Electrode interaction in multichannel cochlear implants with different electrode designs and positions,” in Abstracts of the 2000 annual midwinter meeting, Association for Research in Otolaryngology, 90.

Throckmorton, CS and Collins, LM. (1999). “Investigation of the effects of temporal and spatial interactions on speech-recognition skills in cochlear-implant subjects,” J. Acoust. Soc. Am. 105(2), 861-873.

Townshend, B, Cotter, N, van Compernolle, D, and White, RL (1987). “Pitch perception by cochlear implant subjects,” J. Acoust. Soc. Am. 82, 106-115.

Tye-Murray, N and Tyler, R (1989). “Auditory consonant and word recognition skills of cochlear implant users,” Ear and Hear 10, 292-298.

van den Honert, C and Stypulkowski, PH. (1987). “Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve,” Hearing Res. 29, 195-206.

van Tasell, DJ, Soli, SD, Kirby, VM, and Widin, GP (1987). “Speech waveform envelope cues for consonant recognition,” J. Acoust. Soc. Am. 82(4): 1152-1161.

van Tasell, DJ, Greenfield, DG, Logemann, JJ, and Nelson, DA. (1992). “Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation in cochlear implants,” J. Acoust. Soc. Am. 92(3): 1247-1257.

White, MW, Merzenich, MM, and Gardi, JN. (1984). “Multichannel cochlear implants: Channel interactions and processor design,” Archives of Otolaryngology. 110, 493-501.

White, RL and Van Compernolle, D (1987). “Current spreading and speech-processing strategies for cochlear prostheses,” Ann Otol Rhino Laryng 96(Suppl 128), 22-24.

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

Wilson, BS, Lawson, DT, Zerbi, M, Finley, CC, and Wolford, RD (1995). “New processing strategies in cochlear implantation,” Am. J. Otol. 16(5), 669-675.

Young, ED and Sachs, MB (1979). “Encoding of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers,” J. Acoust. Soc. Am. 66: 1381-403.

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

EDUCATION:

1995-present	Ph.D.Cognition & Neuroscience, University of Texas at Dallas
1996-1998	M.S Audiology, University of Texas at Dallas
1987-1992	B.S. Cognitive Science, University of California, San Diego

RESEARCH EXPERIENCE:

2000	Member of the Editorial Board for *Ear and Hearing*
9/95 - present	Research Assistant, School of Human Development, UT Dallas
9/98 - 10/99	Audiologist, House Ear Institute, Los Angeles, CA
9/93 - 8/95	Laboratory Assistant II, Dept of Psychiatry, UC San Diego
1/92 - 9/93	Research Assistant, Dept. of Cognitive Science, UC San Diego
9/92 - 1/93	Intern, Magnetic Source Imaging Division,VA Med. Center, Albuquerque, NM

CLINICAL EXPERIENCE:

9/98-10/99	Audiologist, House Ear Clinic, Los Angeles, CA
9/97-5/98	Extern, UT Southwestern Medical Center, Dallas, TX
9/96-9/97	Intern, Callier Center for Communication Disorders, Dallas, TX

TEACHING ASSISTANTSHIPS:

Fall 1997	Research Methods
Spring 1997	Behavioral Neuroscience
Summer 1996	Cognitive Psychology

AWARDS

Alcatel Graduate Student Research Fellowship

"Best Student Paper Award in Speech Communication" at the 134th Meeting of the Acoustical Society of America

PUBLICATIONS

Stickney, G. S. and Assmann, P. A. (2001). Acoustic and linguistic factors in the perception of bandpass-filtered speech, Journal of the Acoustical Society of America 109(3): 1157-1165.

Roeser, R. J., Buckley, K.A., and Stickney, G.S. (2000). Pure Tone Tests. In Roeser, R., Valente, M., and Hosford-Dunn, H. (Eds.) Audiology Diagnosis New York: Thieme, pp. 227-251.

CONFERENCE POSTER PRESENTATIONS

Stickney, G. S. and Assmann, P. A. (2000). *Channel interaction and speech processing strategies for cochlear implants*. Journal of the Acoustical Society of America.

Stickney, G.S., Shannon, R., Opie, J., Assmann, P. A. (2000). *Electrode interaction in multichannel cochlear implants with different electrode designs and positions*. Presented at the 23rd mid-winter meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL.

Stickney, G.S. and Shannon, R, (2000). *Simultaneous masking as a measure of electrode interaction in cochlear implant listeners*. Presented at the 6th International Cochlear Implant Conference, Miami Beach, FL.

Stickney, G.S., Shannon, R., Opie, J (1999). *Psychophysical electrode interaction and speech perception in adult cochlear implant listeners with and without the electrode positioning system*. Presented at the 1999 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.

Stickney, G.S. and Assmann, P. A. (1998). *Masking of filtered speech by a single competing voice*. Presented at the 21st mid-winter meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL.

Stickney, G. S. and Assmann, P. A. (1997). *Intelligibility of bandpass-filtered speech*. Journal of the Acoustical Society of America 102(5, Part 2), pg. 3134.

Nava, C., Pineda, J. A. & Stickney, G. (1993*). Human and monkey N400-like responses to faces in a priming paradigm*, Society for Neuroscience Abstracts, 19(2), 1607.

PROFESSIONAL SOCIETIES:

Acoustical Society of America
American Speech-Language-Hearing Association

UNIVERSITY SERVICE

University of Texas Student Advisory Council (1997-1998)
University of Texas at Dallas, Multicultural Association (1996-1998)