Speech Processors for Auditory Prostheses

NIH Contract NO1-DC-92100

QPR #12: October-November-December 2001 and Final Report

Final Progress Report

Effects of stimulation rate on speech recognition with cochlear implants

Submitted by

Robert V. Shannon, Qian-Jie Fu, Monita Chatterjee, John J. Galvin III, Lendra Friesen, Rachel Cruz, John Wygonski, Mark E. Robert
Department of Auditory Implants and Perception
House Ear Institute
2100 W. Third St.
Los Angeles, CA 90057

And

Fan-Gang Zeng
Department of Otolaryngology
University of California, Irvine

7 March 2002

Table of Contents: **PAGE**

Summary of Progress over Last Three Years

Hardware Development:

Over the three years of this contract we have made substantial progress in designing and implementing hardware that provides experimental access to the two major commercial cochlear implants systems: the Nucleus-24 system and the Advanced Bionics Clarion system.

The House Ear Institute Nucleus Research Interface (HEINRI) was developed from the ground up in our lab. This interface convert a stream of bytes from a PC parallel port into the correct code for transmission to the Nucleus-22 and Nucleus-24 implanted receiver/stimulator. This interface allows delivery of any valid biphasic pulse (more than 10 usec/phase in duration and within the amplitude range of the stimulator) to any combination of two electrodes. Speech stimuli can be preprocessed and delivered to a patient through this interface to implement a variety of speech processing strategies. At present we have implemented the SEMA transmission protocol. Which allows delivery of custom stimulation to either the Nucleus-22 of –24 device. A programmable clock can change the output clock rate from 2.5 MHz for the N-22 device to 5.0 MHz for the N-24 device. The Embedded transmission protocol for the N-24 device has not yet been implemented – this will require a completely new version of the real-time DSP code downloaded into the interface unit. With the existing SEMA coding and a 10 usec/phase pulse the present interface is capable of delivering approximately 8 kHz pulses total throughput. This rate allows implementation of 8-channel CIS processors with 1 kHz pulses/sec/electrode.

The Clarion Research Interface 2 (CRI-2) was developed partly under contract from Advanced Bionics Corporation (ABC). This interface will allow complete control of delivery of stimulation to patients with the second-generation CII device from ABC. Analog and pulsatile stimulation are possible, with pulse durations as short as 2 usec/phase. HEI assisted in the hardware development and designed and implemented the low-level DSP software operating system. The CRI-2 will allow presentation of arbitrary pulse or analog waveforms to any combination of electrodes in the CII device at biphasic pulse rates of up to 250 kHz total throughput. The interface is now available to research groups worldwide through ABC.

Rather than design and build our own wearable stimulator we selected to purchase a SPEAR3 wearable research stimulator from the research group in Melbourne, Australia. The SPEAR3 contains a DSP processor in a battery-powered SPRINT package. We only received the SPEAR3 unit in the last quarter of 2001 and have not yet programmed it for experiments with patients.

Stimulation Rate Experiments

We have spent considerable effort to evaluate the effects of stimulation rate on speech recognition performance. Our previous work (Fu and Shannon, 2000) indicated no change in performance as a function of stimulation rate. Work by Blake Wilson and Stefan Brill had shown improved performance at higher stimulation rates and Jay Rubinstein had developed a model of stochastic firing in the electrically stimulated auditory nerve that suggested a new mechanism by which high rates might improve speech recognition. Since this is an important topic for both theoretical an practical reasons, we performed a series of experiments to measure the effect of stimulation rate on speech recognition. Our previous work had only looked at stimulation rates up to 500 pulses/sec/electrode (PPSE) whereas other studies had used pulse rates up to 4500 ppse. Fu and Shannon (2000 showed an improvement in vowel and consonant recognition as the pulse rate was increased from 50 ppse to 150 ppse and then no further improvement from 150 ppse to 500 ppse. In progress reports #2 and #10 of this contract and again in the progress for this quarter we repeated these experiments in listeners with Nucleus-24 and Clarion CII devices in which we could obtain higher stimulation rates. In all 9 listeners we saw no change in performance for stimulation rates between 250 ppse and the maximum rates for each condition and device (up to 2500 ppse). These results are consistent with another study by Vandali et al. (2000) which also observed no change in performance for stimulation rates between 250 and 1500 ppse. Recent results obtained in this quarter with CII listeners show a decrease in performance when stimulation rates exceed 2500 ppse. However, the high stimulation rates in this case were obtained using the QPR stimulation mode, which allows 4 simultaneous pulses to be presented across the electrode array. Thus, it is not possible to know if the decrease in performance at high rates was due to rate or to electrical field interactions due to the simultaneous presentation.

Spectral Asynchrony

When sound is processed by an implant speech processor rather than the normal cochlear mechanics it is possible that the normal timing relationships across frequency are disrupted. When a burst of energy is produced in speech this will produce a burst of energy in all frequency regions, delayed by the group delay of the band-pass filtering inherent in the cochlea. It is possible that the central auditory system uses such cross-frequency synchrony to mark brief temporal events, ot to synchronize temporal events across frequency. To measure the effect of cross-frequency temporal synchrony, we systematically introduced asynchrony across processing filter bands and measured speech recognition. We presented the results of this study in Quarterly Report #6 and published the results in Fu et al. (2000). The results showed that both CI and NH listeners are remarkably resistant to the effects of cross-frequency asynchrony, as long as the spectral information was intact. With original speech, time delays of up to 240 msec across the frequency range only resulted in miminal reduction in recognition. However, this effect was larger when spectral resolution was poor, as in cochlear implants. When spectral information was represented by 16 modulated noise bands, performance fell to 60% correct with a cross-frequency asynchrony of only 160 msec. When spectral resolution was reduced to 4 channels, both NH and CI listeners showed significant drop in speech recognition when only 40 msec delay was introduced across the frequency range. This result shows a clear trade-off between spectral resolution and the ability to tolerate spectral asynchrony. However, even with poor spectral resolution this result shows that listeners are relatively resistant to temporal asynchrony of as much as 20 msec. This values is considerably larger than any group delay that would be introduced across frequency by a speech processor and more than the 5 msec delay that occurs across frequency in the normal cochlea.

Number of Channels

Performance with cochlear implants is strongly tied to the number of channels used in the processor for the representation of spectral information. Previous research has shown that good speech recognition can be achieved with as few a 4 spectral channels, but more channels are necessary for difficult speech materials, difficult listening conditions, and competing noise. We measured speech recognition on a variety of speech materials in conditions of noise as we varied the number of stimulated electrodes. We provided preliminary data from this study in quarterly reports #1 and #2 and the final results were presented in QPR#5. For normal hearing (NH) listeners speech recognition improves as the number of spectral channels is increased. The largest improvement is observed as the number of channels increases from 1 to 8, with a smaller increase as the number of channels is increased above 8. CI listeners however, appear to asymptote in performance at about 8 channels. An increase in the number of channels above 8 provides no improvement in performance, even in noise (Friesen et al., 2001). It is unclear why CI listeners are unable to utilize the full number of spectral channels of information presented to their device. We speculate that errors in mapping frequency to electrodes may cause some of the problem (see frequency-to-place mapping section below). Another possible factor limiting performance is the interaction between channels, and measures of channel interaction were presented in QPR#7.

Amplitude Mapping Effects

Our objective was to study how to optimally set the acoustic dynamic range and map it into the electric dynamic range in cochlear-implant listeners. In QPRs #4 and #9 we presented empirical data on acoustic dynamic range, electric dynamic range, and electrode interactions. We also reported how these dynamic ranges and electrode interactions affect speech recognition in Clarion and Nucleus implant users. Acoustic analyses of phoneme tokens produced by 5-female and 5-male talkers showed a 50 dB speech dynamic range, much wider than the commonly assumed 30 dB range. Psychophysical data collected over a large set of parameters including monopolar and bipolar modes, sinusoidal and pulsatile waveforms, stimulus frequency, and pulse duration showed that electric dynamic range rarely exceeds 30 dB. Modeling and experimental data indicated that electrode interactions should be taken into account when fitting speech processors of modern multi-electrode cochlear implants. Direct electrical field summation across electrodes reduces the effective number of independent channels, while loudness summation reduces the effective dynamic range under multiple electrode stimulation. Corresponding speech recognition experiments also yielded interesting results. Peak performance was achieved when the acoustic dynamic range was set at 50 dB, consistent with the acoustic analysis results. Reducing electric dynamic range produced little effect on Nucleus users of the SPEAK strategy but decreased speech recognition in Clarion users of the CIS and SAS strategies. When electrode interactions were taken into account in a new “speech-adjusted” fitting strategy, noticeable improvement in speech recognition was observed in two Clarion users. The data indicate that both acoustic and electric dynamic ranges need to be optimized in cochlear implants and different optimizations may be needed for different processing strategies. Furthermore, the results suggest that electrode interactions must be taken into account in speech processor fittings and new fitting protocols should be developed to use wide-band, dynamic speech to fit cochlear implants under realistic listening situations. These new fitting protocols may be particularly useful in children who use cochlear implants.

Frequency-Place Mapping

Previous research has shown that a distortion in the mapping of frequency information to cochlear place can result in major reduction in speech recognition performance. Such frequency-place mapping can occur in cochlear implants due to differences in electrode placement, and do occur as a result of the normal speech processor frequency-to-electrode assignments. The typical cochlear implant electrode array is positioned approximately 25 mm inside the round window in the scala tympani. The active electrode contacts occupy a cochlear region ranging from 9 to 25 mm from the round window, a frequency region that would normally respond to acoustic frequencies of 500-5000 Hz according to Greenwood's (1990) formula. However, most cochlear implant speech processors divide the entire spectrum from 100-10000 Hz into bands and present the information from the bands to the electrodes. This broad frequency range would normally occupy cochlear locations from 3 to 30 mm from the round window – a much larger range than that occupied by the CI electrode array. This results in a two octave compression of frequency-place; frequency information that would normally be represented across a distance of 25-30 mm of the cochlea are being presented to electrodes that span only a 16 mm range.

We measured vowel, consonant and sentence recognition in NH listeners with noise-band speech processors with 4, 8 or 16 channels. In each of these conditions we varied the frequency-place mapping from 5 mm expansion of frequency-to-place on each end of the spectral range, to frequency-place match, to 5 mm compression of frequency-to-place on each end of the spectral range. This last condition simulates the normal frequency-to-electrode assignment of a cochlear implant.

The results (manuscript presented in Quarterly report #8 and submitted for publication, Baskent and Shannon, submitted to JASA) show a 20% decrease in performance with 5 mm compression and a larger decrease with 5 mm expansion. Performance was usually best when the frequency range matched the range of noise carrier bands simulating the electrode location. There was a trade-off between compression and simulated electrode location. If the noise bands simulating electrode location were shifted basally, like an electrode that was not fully inserted, then the best performance was observed with some degree of frequency-place compression. In this case a strict matching of frequency to place results in the elimination of too much critical low-frequency information. Thus, some frequency-place distortion that includes those low frequencies is better that a strict frequency-place match that loses information.

Another type of frequency-place distortion is what we have termed a "hole" in hearing – resulting from a loss in neurons in a local tonotopic range. An electrode in that region would have no local neurons to stimulate. If the current level on that electrode were increased the current field would eventualy spread out and activate neurons in an adjacent tonotopic region. This would have a localized distortion effect in the frequency-place map: frequency information that would normally be presented to one tonotopic area would be represented in an adjacent tonotopic area. We measured speech recognition in NH listeners with 20-channel noise-band processors to simulate the effects of such a “hole” in the tonotopic representation and the resulting distortion in the frequency-place map. Speech was analyzed into 20 frequency bands and the envelope was extracted. Twenty noise bands were used to simulate 20 electrode locations. “Holes” were simulated by eliminating several noise carrier bands to make “holes” in the tonotopic representation from 1 to 6 mm in extent. Holes were simulated in the apical, middle and basal regions of the cochlea. Three manipulations were done in an attempt to rescue the envelope information that would normally go to the hole region: the envelope information from the hole was assigned either to the apical edge of the hole, the basal edge of the hole, or even split to the two sides of the hole. The results, presented in Quarterly Reports #1, 2, and 3 (manuscript published in JARO, Shannon et al., 2002), show that even large holes in the basal and middle spectral regions have only a small effect on speech recognition. Holes in the apical region can have large detrimental effects on speech recognition. The three types of reassignment conditions were not successful in rescuing the information from the hole region – all three reassignment conditions produced speech recognition equivalent to simply dropping the information. So, at least with no time to adapt to the new map, the speech information from the hole region was lost, whether the information was simply deleted or moved to an adjacent tonotopic region. It is possible that this distorted pattern of information might be usable after a period of relearning. To investigate this possibility we embarked on a series of experiments to investigate the effects of training and learning on adaptation to an altered frequency-place map.

Effects of Long-Term Learning

One question about the effects of frequency-place distortion is whether listeners will simply and quickly accommodate to such distortions. To answer this question we shifted the frequency-to-electrode map down in frequency by one octave in three CI listeners. Listeners wore the experimental map as their only map for three full months. Performance was measured a weekly intervals to monitor any changes in performance over time as they accommodated to the new map. In most speech measures we observed a rapid improvement in performance in the first few weeks and then little improvement over the rest of the three-month experimental period. There appeared to be two components to long-term learning – a relatively rapid accommodation that allowed performance to recover somewhat in the first few weeks, and a longer-term accommodation that was far from complete at the end of three months. Either that second, longer accommodation has a very long time constant of recovery, or the subject would never have accommodated to the one octave apical shift in the pattern of tonotopic information. Following the experiment the listeners’ original frequency-to-electrode maps were restored and performance was again assessed. Following three months with a shifted map, their speech recognition was identical to their original performance level. Thus, whatever accommodation the listeners made during the three month trial with a shifted map did not interfere with their understanding with their original map. This again suggests that the time constant for adapting to a shifted frequency-place map is quite long. These results were described in Quarterly Reports #1 and #2, and in a manuscript that is in press (Fu and Galvin, 2002)

A second experiment (Quarterly Report #11) looked at accommodation to a frequency-place shift by NH listeners with noise-band processors. NH listeners were given training with feedback on a 20 channel noise-band processor fo 5 consecutive days in a laboratory setting. They were trained on a processor that was shifted approximately 5 mm basally. After two hours of training on five successive days their speech recognition was assessed on processors with a variety of frequency-place shifts. It was observed that listeners improved significantly on the trained condition and that the training did not generalize, i.e., that their performance did not improve on conditions with other frequency-place shifts. Although an improvement in performance was observed even after a modest amount of training, performance was still considerably poorer that when frequency information was presented to the normal acoustic place.

Experimental Report:

Effects of stimulation rate on speech recognition with cochlear implants

INTRODUCTION

Cochlear implants (CIs) have significantly improved since the early single-channel models. New devices have 8-22 electrodes and are capable of stimulating pulse rates in excess of 1000 pulses/second/electrode (ppse). Multiple electrodes have resulted in significant improvements in speech recognition over single-channel models. New speech processor designs and electrode designs have also resulted in increased performance. One potential limitation of multiple electrodes is the overlap of current fields between adjacent electrodes. To avoid the overlap of current fields from adjacent electrodes in a multi-electrode array, most CI stimulation strategies use temporally interleaved biphasic current pulses. For a large number of electrodes, temporally interleaved stimulation necessarily reduces the overall rate of stimulation of each electrode. For example, in a 20-electrode device with 100 µs/phase biphasic current pulses the maximum theoretical rate of stimulation would be 250 ppse. A stimulation rate of 250 ppse may not be high enough to convey temporal changes relevant to speech features. According to the sampling theorem a stimulation rate of 250 ppse would only be able to represent temporal changes up to 125 Hz. To achieve higher stimulation rates requires shorter pulses, which in turn require higher pulse amplitudes to achieve the same perceptual effect. Higher pulse amplitudes would produce a wider spread of the instantaneous current field, possibly increasing the overlap of the neural populations excited by each electrode. Thus, there may be a trade-off inherent in the use of high stimulation rates in cochlear implants. Temporally interleaved pulses are used to avoid current field interactions, but temporal interleaving limits the ability to represent the fine temporal structure of the stimulus.

And to achieve higher stimulation rates requires shorter pulses and higher amplitudes, which may create more overlap in the neural populations activated.

Perceptual Effects of Stimulation Rate

Data on the effect of different stimulation rates on speech recognition has so far been inconclusive. Brill et al (1997, 1998) conducted three experiments with the Med-El COMBI 40+ device. In the first experiment one subject showed an increase in consonant recognition scores when the pulse rate was increased from 625 to 1515 ppse. In the second experiment, monosyllabic word recognition scores peaked at the rates of 1515 ppse for three subjects and at 3030 ppse for two other subjects with 6-channel processors when the rate was varied between 400 and 3030 ppse per electrode. In the third experiment, monosyllabic word recognition was tested while the stimulation rate was varied between 800 and 4545 ppse in 4-channel processors. Two subjects showed an increase in scores with an increase in rate up to 4545 ppse, the highest rate tested, one subject showed a peak in scores at 1515 ppse, and one other subject showed asymptotic performance at 3030 ppse.

Lawson *et al.* (1996) measured consonant recognition in cochlear implant listeners with six-channel CIS processors, utilizing rates of 250, 833, and 2525 ppse. Scores were similar for the three rates.

Loizou and Poroy (1999) measured monosyllabic word and phoneme recognition in 6 cochlear implant listeners using the CIS strategy with pulse rates of 400, 800, 1400, and 2100 ppse. An improvement in scores was noted for consonant and monosyllabic word recognition when rates were increased from 400 to 800 and from 800 to 2100 ppse. Vowel scores remained relatively unchanged as rates were varied.

Fu and Shannon (2000) measured vowel and consonant recognition in six Nucleus-22 cochlear implant listeners while varying pulse rate between 50 to 500 ppse, using an experimental 4-channel CIS speech processing strategy. Performance improved as pulse rate increased from 50 ppse to 150 ppse, but there was no improvement in performance for stimulation rates between 150 and 500 ppse.

Kiefer et al. (2000) measured phoneme and word recognition in 13 subjects with the Med-El COMBI-40 and COMBI-40+ cochlear implants as a function of stimulation pulse rate. They found a small, but significant improvement in performance at when the stimulation rate was increased from 600 ppse to 1515-1731 ppse. No significant improvement in performance was observed as pulse rate was increased from 1515 ppse to as high as 4545 ppse.

Vandali et al. (2000) measured monosyllable word recognition in quiet and sentence recognition in noise in six Nucleus 24 listeners at stimulation rates of 250, 807, and 1615 ppse with the ACE (8/22) speech processing strategy. They found no significant difference in performance between 250 and 807 ppse, and significantly poorer performance at 1615 ppse on some tests. Although there were no significant differences in performance between 250 and 807 ppse, individual patients preferred the sound quality at one of the two rates over the other. No patient preferred the sound quality at 1615 ppse.

In sum, some studies show an improvement in speech recognition as the stimulation pulse rate is increased and other do not. Some studies show an improvement in speech recognition as rate is increased for some subjects and not for others. And some studies showed a decrease in speech recognition at very high stimulation rates.

Physiological Effects of Stimulation Rate

High stimulation rates may be important for reasons beyond the ability to represent fine temporal features of the stimulus: high stimulation rates may produce more normal temporal patterns of nerve activity. Wilson *et al.* (1997a) have demonstrated that electric stimulation causes an alternation in the neural response due to a complex interaction between electrical activation and refractory properties. Electrical stimulation fires all local neurons synchronously, so they all enter the recovery cycle at the same time. If the time till the next stimulation pulse is long enough they will all be ready to fire again. If the time to the next pulse is short none of the neurons will be recovered sufficiently to be able to fire on the next pulse. These competing processes lead to a distinctive alternation pattern of neural firing for medium stimulation rates. The neural alternation, which is a form of signal aliasing, only disappears when the stimulation rate is very high – above 2-3000 ppse.

Visher *et al.* (1997) measured electrically-evoked auditory brainstem responses (EABRs) in implanted rats stimulated with rates from 100 ppse to 1500 ppse. As the stimulating pulse rate was increased, the EABR wave 1 amplitude dropped to 15% of its single-pulse value by 1500 ppse. They speculated that a lower ABR amplitude represented less synchrony in the auditory nerve, consistent with Wilson et al.'s results. Both Wilson and Visher speculate that this more normal response pattern might translate into better speech understanding.

Another potential benefit of high stimulation rates is the restoration of normal stochastic properties in the neural response (Rubinstein *et al.*, 1999). In acoustic stimulation each neuron fires probabalistically on each cycle of a stimulating tone. The timing of the nerve discharge is time-locked to the phase of the stimulating tone up to about 5 kHz. In electric stimulation, neurons fire in a more deterministic manner. Once the electric current level exceeds the threshold for activation, all neurons fire synchronously. At low stimulation rates neurons will fire on every stimulation pulse. Only at high stimulation rates will neurons fire probabilistically to electric stimulation, because the probabilistic recovery from the absolute refractory period de-synchronizes the individual neurons. On each individual pulse only a subset of the neurons are recovered from their previous action potential sufficiently that they are available to fire to the present pulse. Rubinstein et al. speculate that this stochastic firing may improve the ability of implant listeners to utilize temporal information in speech as so improve speech recognition.

In sum, high electric stimulation rates should provide improved representation of temporal information in speech, and may produce more natural stochastic patterns of discharge in the overall neural ensemble.

The present study measured phoneme, word and sentence recognition using the same testing protocol in a nine subjects with two different implant devices, using stimulation rates available with the clinical fitting systems.

METHODS

Listeners

Five adults (18 years and older) using the Clarion C1 cochlear implant, 3 adults with the Clarion CII, and 4 adults with the Nucleus 24 implant, each having at least six months experience, participated in this study. All were postlingually deafened and native speakers of American English. General demographic information for the 9 subjects is presented in Table 1. All Clarion users had eight electrode pairs available for use and all Nucleus-24 listeners had 22 available electrodes.

Speech materials

Speech perception tests were all presented without lip-reading (sound only). The tests consisted of medial vowel and consonant discrimination, monosyllable word recognition and sentence recognition.

Vowel stimuli were taken from materials recorded by Hillenbrand et al. (1995) and were presented to the listeners with custom software (Robert, 1997). Ten presentations (5 male and 5 female talkers) each of twelve medial vowels in a h/V/d context (/i ɪ ɛ æ u ʊ ɑ ʌ ɔ ɚ o e/) presented in a /h/-vowel-/d/ context (heed, hid, head, had, who'd, hood, hod, hud, hawed, heard, hoed, hayed). Chance level on this test was 8.33% correct and the 95% confidence level was 11.8% correct.

Consonant stimuli (3 male and 3 female talkers in a /a/C/a/ context) were taken from materials created by Turner et al. (1995, 1999) and Fu et al. (1998). Consonant confusion matrices were compiled from 12 presentations of each of 14 medial consonants /b d g p t k m n f s ʃ v z θ/, presented in an /a/-consonant-/a/ context. Tokens were presented in random order by custom software (Robert, 1997) and the confusion matrices were analyzed for information received on the production based categories of voicing, manner, and place of articulation (Miller and Nicely, 1955). Chance performance level for this test was 7.14% correct, and the 95% confidence level was 10% correct.

Consonant stimuli for the Clarion CII listeners were 20 consonants in /aCa/ context taken from the materials recorded by Shannon et al. (1999). These materials consist of natural tokens spoken by five male and five female talkers.

The CNC Word Test from the Minimum Speech Test Battery for Adult Cochlear Implant Users CD was used to evaluate open-set phoneme and word recognition (House Ear Institute and Cochlear Corporation, 1996). The CD contains 10 lists of 50 monosyllabic words containing 150 phonemes. Listener responses were scored separately for words and phonemes correctly identified.

Recognition of words in sentences was measured using the Hearing in Noise Test (HINT) sentences from the Minimum Speech Test Battery for Adult Cochlear Implant Users CD (House Ear Institute and Cochlear Corporation, 1996). For each condition, data was collected for 10 sentences of varying lengths from each listener. The sentences were of easy-to-moderate difficulty, presented with no context and no feedback, and no sentences were repeated to an individual listener. For Clarion CII listeners sentence recognition was evaluated with IEEE sentence materials, which are moderately difficult sentences with only a minor amount of contextual information.

Experimental speech processor conditions

There were a total of 7 experimental speech processors evaluated for each Clarion C1 listener, 22 processors for each Clarion CII listener, and 12 processors for each Nucleus-24 listener. Testing was performed immediately after listeners received them (no practice).

Clarion C1 listeners were tested with 4-electrode and 8-electrode processors, each at a variety of stimulation rates. With the Clarion CIS speech processing strategy, 8 frequency bands are normally directed to 8 electrode pairs (Clarion by Advanced Bionics, 1998). All testing with Clarion C1 devices was done in bipolar stimulation mode. All Clarion C1 patients had the “enhanced” Bipolar Electrode with the positioning system (EPS). With a reduction in the number of electrode pairs to 4, the total frequency range remained the same, but the range for each electrode was broadened. Stimulation rates were 200, 400, and 800 ppse for the 8-electrode processors and 200, 400, 800, and 1600 ppse for the 4-electrode processor. Electrode pairs 1, 3, 5, and 7 were activated in the 4-electrode processors. All processors and electrode conditions used a pulse duration of 75 μs/phase.

Clarion CII listeners were tested with 4, 8, 12, and 16 electrode processors at a variety of stimulation rates. In all conditions pulses were 11 μs/phase biphasic pulses and the stimulation mode was monopolar. All Clarion CII listeners had the Hi-Focus electrode array and the electrode positioning system (EPS). The electrode array in the Clarion CII device is a linear array of contacts, rather than the paired bipolar sets of electrodes of previous models. The CII contains 16 independent current sources that allow monopolar stimulation of up to 16 independent channels. Experimental speech processor software (Bionic Ear Programming System: BEPS) allowed selection of electrodes and stimulation rates for CIS processors. Sixteen electrode processors used all available electrodes. Twelve-electrode processors eliminated every fourth electrode, so that the remaining electrodes were 1-2-3-5-6-7-9-10-11-13-14-15. The same frequency range was used for all processors, but the range was divided into fewer number of frequency bands when fewer electrodes were used. Eight-electrode processors used the odd numbered electrodes, and four-electrode processors used electrodes 1-5-9-13. The BEPS software allows stimulation rates up to 11,363 ppse for 4-channel CIS processors, with all pulses strictly interleaved. Higher rates can be achieved by allowing some pulses to be presented simultaneously. In the present experiment a few very high rate conditions were implemented using the QPS stimulating strategy. In QPS strategy the entire electrode array is divided into four quadrants. The CIS strategy is implemented within each quadrant, with four pulses presented simultaneously on electrodes from each quadrant. Since these processors all used monopolar electrode mode, simultaneous stimulation opens the possibility of electrical current field interaction between the simultaneously stimulated electrodes. With the QPS strategy stimulation rates up to 15,475 ppse are possible in 4-channel processor.

The Nucleus-24 device with the ACE processing strategy allows for the determination of the sites of stimulation (maximum 22), the number of maxima during each stimulation frame (maximum 20), and the stimulation rate per channel (maximum 14,400 Hz overall). During programming, the clinical fitting system automatically assigns frequency information to electrodes based on the processing strategy selected and number of active channels. The default frequency allocation tables were used for all experimental processors. Four conditions were selected using 4, 8, 12, or 16 electrode pairs [electrodes used in the 4-electrode processor: 4, 10, 16, 22; 8-electrode processor: 4, 7, 10, 13, 16, 18, 20, 22, 12-electrode processor: 4, 6, 8, 10, 12, 14, 15, 16, 18, 19, 20, 22, and 16-electrode processor: 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 22]. Processors were programmed with the fastest and slowest rates allowable in the clinical system, plus at least one intermediate rate. Stimulation rates ranged from 200 Hz to 2400 ppse, depending on the number of electrodes used. Stimulation rates were 250, 900, and 2400 ppse for the 4-electrode condition, 250, 900, and 1800 ppse for the 8-electrode condition, 250, 900, and 1200 ppse for the 12-electrode condition, and 250, 500, and 900 ppse for the16-electrode condition. The pulse duration for all conditions was 25 µs/phase.

Procedure

During all testing the listener was seated one meter in front of a loudspeaker (Grason-Stadler audio monitors) in a sound treated room (IAC). The presentation level was 65 dB SPL for all speech perception testing, as measured by a B&K one inch microphone (Model #4144) at the location of the listener’s head. All speech materials were prerecorded. A computer with a sound card (Turtle Beach Fiji), CD player, and a GSI audiometer (Model 16) was used to present the test items. The GSI 16 audiometer generated the speech-shaped noise used during the vowel, consonant and word tests for the implant listeners. The CD utilized for presenting the HINT sentence materials provided the speech-shaped noise during that test.

Threshold (T) and most comfortable (M) loudness/maximum comfort (C) levels were measured separately for each rate condition. The experimental processors were presented to each listener in random order. The battery of speech tests was administered to each listener immediately after they were given the experimental processor (no practice).

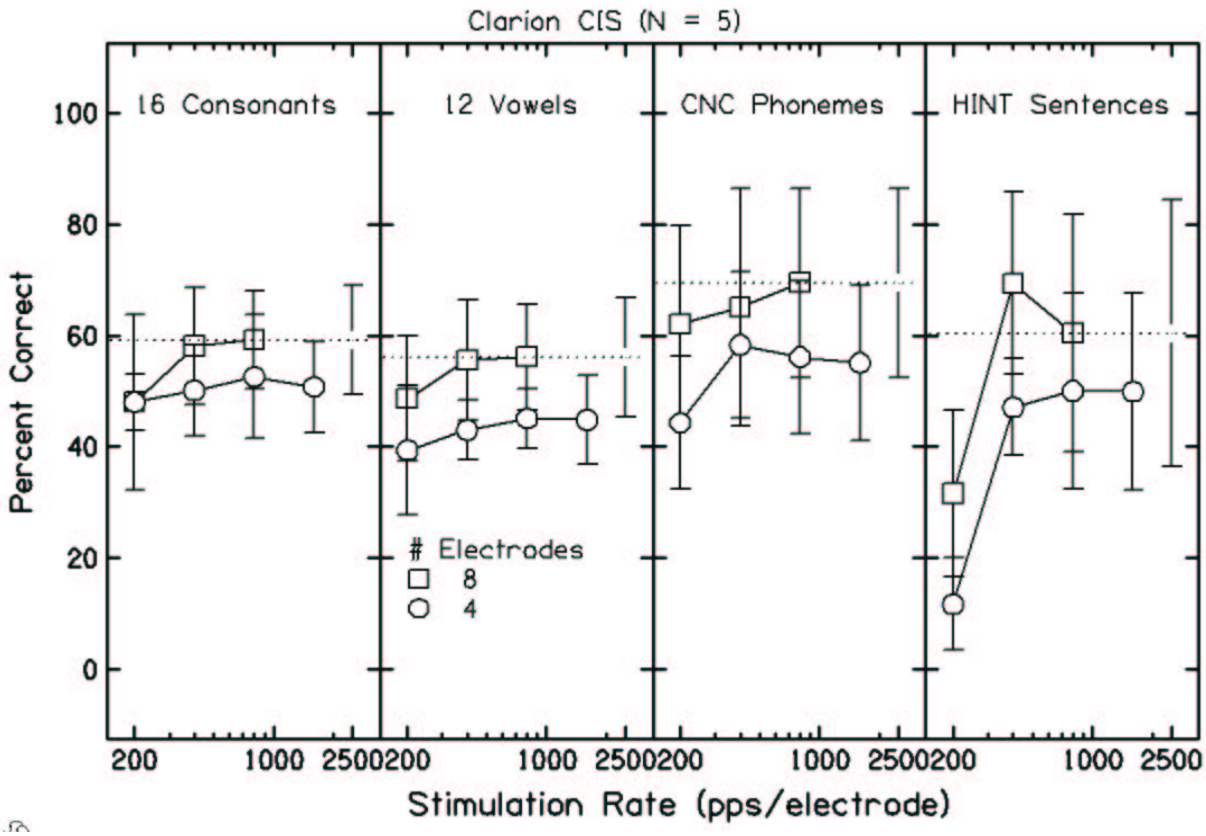
For the Clarion device electrical thresholds (T) and most comfortable loudness (M) levels were obtained using the SCLIN for Windows software, Clinician’s Programming Interface (CPI), and power supply with a personal computer. The Input Dynamic Range was set to -60 dB SL for all conditions. All other parameters were set similarly to the listener’s original processor. In the CIS processing strategy, threshold levels were estimated by a standard clinical bracketing procedure. Initially, all the electrodes were screened for threshold level and the patient was instructed to identify when they first heard the sound. Then, going back to the first electrode, one to five pulse bursts were presented and the listener was instructed to count the number heard. The T level used in the processor was the level at which the listener counted the number of bursts correctly 50% of the time. To obtain M levels the experimenter increased the electrical level until the listener felt the loudness was at the most comfortable loudness level (the level where they heard the sound at a normal conversational level and could listen to it for a long time without discomfort). Adjacent electrodes were balanced for loudness at M level for each electrode.

With the Nucleus-24 CIS processing strategy, electrical thresholds (T) and maximum comfortable loudness (C) levels were obtained using the WinDPS software, PCI, and computer. A SPRINT processor was used for all the testing. Original processors for all subjects contained the SPEAK strategy. Subjects were switched to ACE for this study. Threshold levels were estimated by a standard clinical bracketing procedure. Initially, all the electrodes were screened for threshold level and the patient was instructed to identify when they first heard the sound. Then, going back to the first electrode, one to five pulse bursts were presented and the listener was instructed to count the number heard. The T level used in the processor was the level at which the listener counted the number of bursts correctly 100% of the time. To obtain C levels the experimenter increased the electrical level until the listener felt the loudness was at the maximum comfortable loudness level (the level where they could listen to it for a long time without discomfort). Adjacent electrodes were balanced for loudness at C level for each electrode. The same number of maxima and electrodes were always selected when programming the processor.

RESULTS

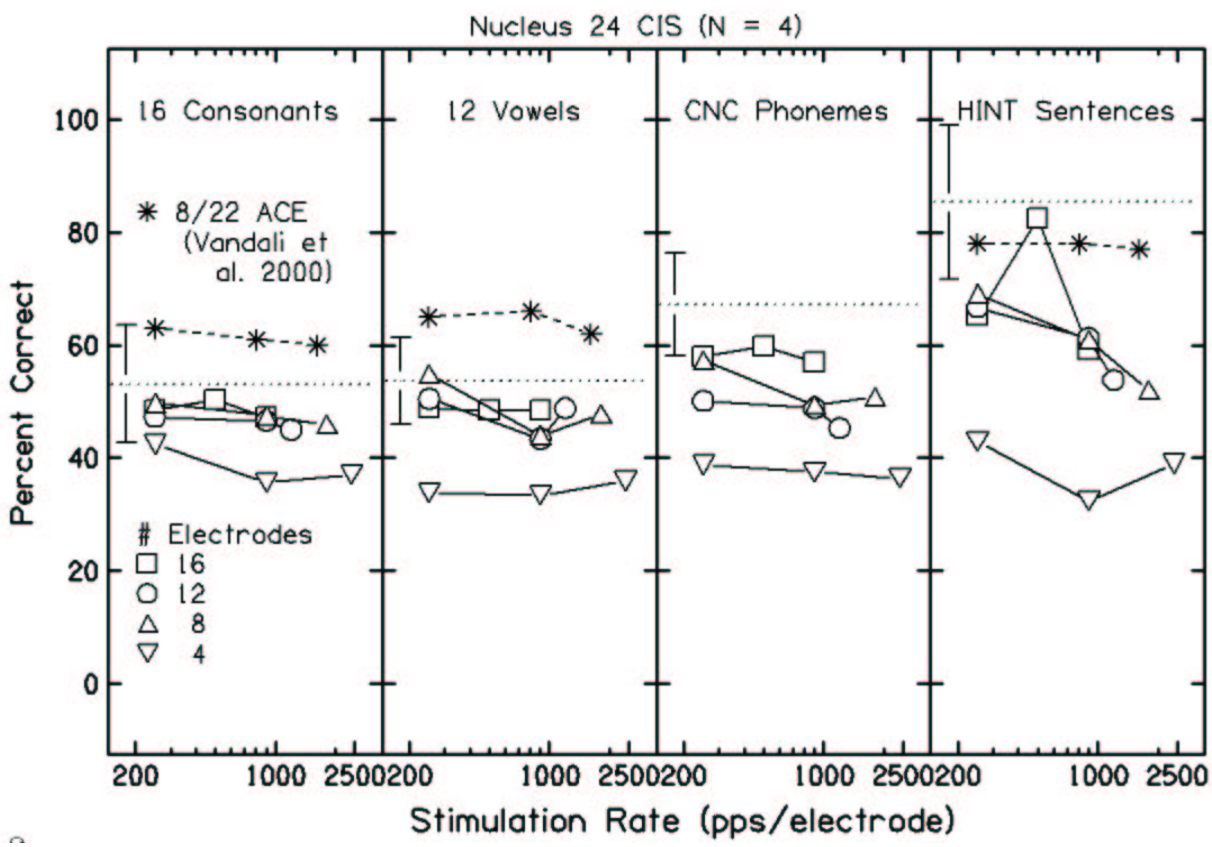
All results are presented as percent correct as a function of stimulation rate in ppse. The average performance with the subjects' clinical device is presented as a dashed horizontal line in each figure. To reduce clutter in the figures error bars are not plotted on all experimental conditions. Error bars indicating one standard deviation on the same test measures for the subject's own processor are shown on the edge of each panel. In most cases the standard deviation for the experimental conditions was similar to or smaller than those shown for the original processor condition.

With the Clarion device, scores for all speech tests tended to increase from 200 to 400 ppse and then plateau as rate was further increased, for both 4- and 8-electrode processors. Repeated measures ANOVAs revealed no significant differences among the scores for 4-and 8-electrode processors for



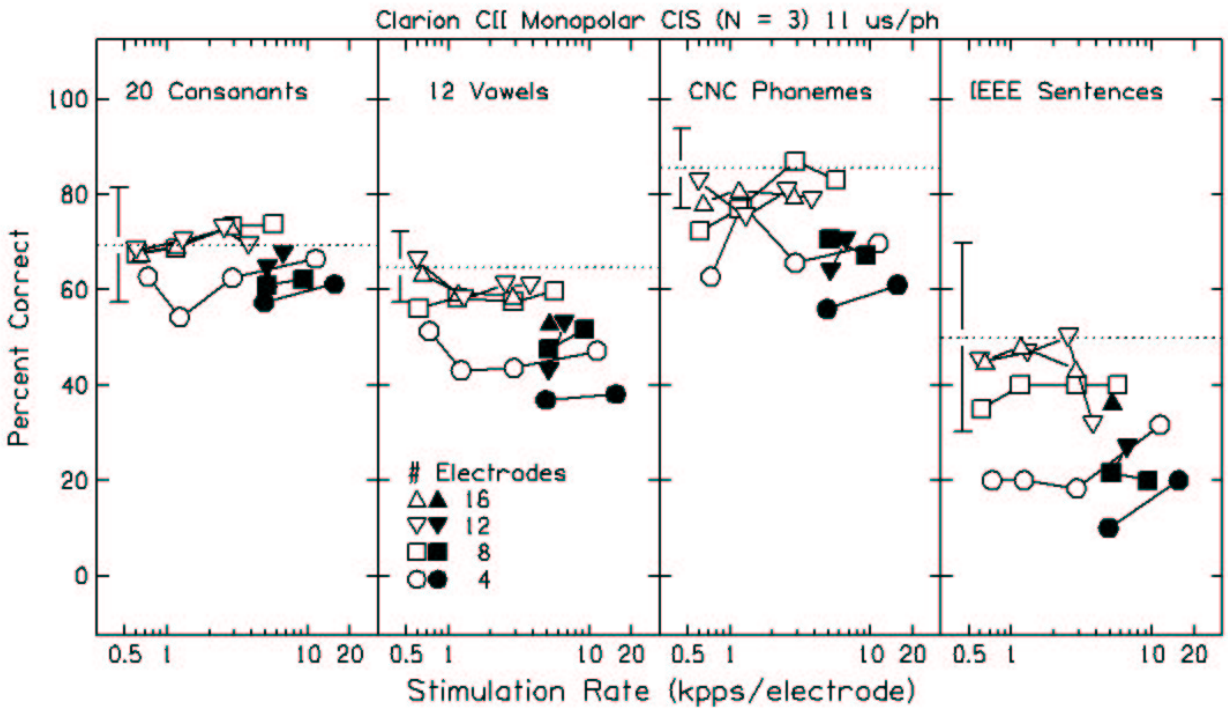
consonants ($F = 0.497$, $p > 0.05$), vowels ($F = 3.263$, $p > 0.05$), words ($F = 2.347$, $p > 0.05$), and sentences ($F = 3.933$, $p > 0.05$). However, there was a statistically significant difference among scores as a function of stimulation rate for consonants ($F = 5.938$, $p < 0.05$), vowels ($F = 4.305$, $p < 0.05$), and sentences ($F = 28.749$, $p < 0.05$). No significant differences were noted with the word scores, probably due to the large variability in scores across listeners. A repeated measures ANOVA was performed with only the 4-channel processors to determine whether there was a statistically significant difference between the 1600 ppse rate and the other rates. The HINT sentence test was the only test for which there was a significant difference in score with a change in rate (consonants: $F = 0.768$, $p > 0.05$; vowels: $F = 0.862$, $p > 0.05$; words: $F = 3.051$, $p > 0.05$; sentences: $F = 11.130$, $p < 0.05$). This difference is probably due to the low scores obtained with the 200-ppse rate programs as compared to the higher scores achieved with all of the faster rate programs.

For the Nucleus-24 cochlear implant listeners, a repeated measures ANOVA with the different channel-numbered processors (between subjects) was performed versus the two rates of 250 and 900 ppse (within-subjects). When the 4-electrode processor was included in this analysis, there was a significant difference among scores for the 4-, 8-, 12-, and 16-channel processors for vowels and sentences



(consonants: $F= 1.813$, $p > 0.05$; vowels: $F= 8.740$, $p < 0.05$; words: $F= 2.931$, $p > 0.05$; and sentences: $F= 5.110$, $p < 0.05$). However, when the 4-channel processor scores were left out of the analysis, none of the differences in score between processors were significant (consonants: $F= 0.058$, $p > 0.05$; vowels: $F= 0.194$, $p > 0.05$; words: $F= 1.108$, $p > 0.05$; and sentences: $F= 0.050$, $p > 0.05$). For the within-subject comparison of scores obtained with different rates, only the vowel test revealed a statistically significant difference in score (consonants: $F= 1.297$, $p > 0.05$; vowels: $F= 6.988$, $p < 0.05$; words: $F= 0.567$, $p > 0.05$; and sentences: $F= 0.692$, $p > 0.05$). It should be noted that the listeners scored higher with the 250-ppse processor than the 900-ppse processor, which is somewhat contrary to previously published results.

Figure 3 presents the average results from the three Clarion CII listeners. Stimulation rates ranged from 595 to 11363 Hz in interleaved stimulation mode and from 5128 to 15475 Hz in QPS mode. No clear pattern of increase or decrease in speech recognition is seen as a function of stimulation rate. Four-channel stimulation results in lower performance than 8-, 12-, and 16-channel stimulation, but there were no significant differences between performance on 8-, 12-, or 16-channel processors. QPS stimulation mode (filled symbols) resulted in lower performance than CIS stimulation modes at the same stimulation rates. QPS stimulation mode resulted in



lower performance than CIS modes even when the QPS was at higher stimulation rates.

Consonant recognition data were analyzed for information received on the traditional production-based categories of voicing, manner and place of articulation (Miller and Nicely, 1955). As expected, more place information was received as the number of electrodes increased from 4 to 8, but no significant increase was observed as the number of electrodes was increased from 8 to 16. Reception of voicing and manner increased as the pulse rate was increased from 200 to 400 for Clarion C1 listeners, but no other change was observed as a function of stimulation rate for any other condition. Table 2 presents the average information percent received on consonant voicing, manner and place for each condition, averaged across all stimulation rates above 400 ppse. Note that the percentage of received information in each for voicing and manner is similar between the Clarion C1 and Nucleus-24 devices. Nucleus-24 listeners appear to receive slightly less information on place than Clarion C1 users, possibly because the experimental speech processing strategies are CIS and they are used to SPEAK or ACE processing strategies. Clarion CII results show considerably higher scores on all feature categories, particularly on voicing and manner. Due to the small sample size it is not possible to draw any conclusions about the relative performance of different devices.

DISCUSSION

The results of this study support previous studies in showing little effect of stimulation rate on speech recognition performance in cochlear implants. Regardless of the implant device, electrode design, or stimulation mode, subjects did not show a significant change in speech recognition for stimulation rates above 400 ppse. Stimulation rates of 400 ppse and higher should provide useful information about temporal features up to 200 Hz, including prosodic envelope features and voicing fundamental frequency. Physiological data (Wilson et al., 1999; Rubinstein et al., 1999; Visher et al., 1999) implies that stimulation rates higher than 400 Hz are necessary to transmit temporal information up to 200 Hz because of “aliasing” and other complications of recovery from refraction. However, analysis of speech consonant features received in the present experiment showed no change in envelope or voicing information received for stimulation rates above 400 ppse. This suggests that higher rates do not provide improved temporal envelope information or improved access to periodicity information.

<u>Possible device or signal processing strategy effects?</u>

The principal studies showing improvement in speech recognition with increased stimulation rate across all test materials and subjects are with the CIS processing strategy in the Med-El COMBI 40+ device (Brill et al., 1998 a,b). None of the other studies exploring this effect show a consistent improvement in speech recognition with an increased stimulation rate across all subjects and test materials (Lawson et al., 1996; Loizou and Poroy, 1999; Fu and Shannon, 2000; Kiefer et al., 2000; Vandali et al., 2000). It seems unlikely that the difference in the results could be due to differences in the implant devices. All modern cochlear implant devices are multichannel devices that are capable of presenting stimulation are rates exceeding 1000 ppse for a 16-channel processor. The Brill et al. results were obtained at much stimulation rates that were used in most of the other studies. However, the present results (Figure 3) show no clear improvement in speech recognition from stimulation rates of 1000 ppse up to rates as high as those by Brill and colleagues. The present results also use the CIS processing strategy.

<u>Learning Effects</u>

Speech recognition with experimental speech processors were generally equal to or lower than the speech recognition with the listeners’ own implant speech processing strategy. With the Clarion device 8-channel processors generally allowed a similar level of performance to that obtained with their clinical speech processor, probably because both experimental and clinical processors used the same strategy (CIS) and same number of electrodes (8). Nucleus-24 listeners showed slightly lower speech recognition for the experimental processors than for their regular clinical processor. This may be due to the difference in speech processing strategy (CIS vs SPEAK/ACE) or to a difference in the number of electrodes (16 vs 21). However, since the present results show no significant difference in results between 8, 12, and 16 electrode processors, it seems unlikely that the difference between 16 and 21 electrodes would produce a significant improvement. This result is in line with previous studies that showed no difference in speech recognition between 7, 10, and 20-electrode processors (Fishman et al., 1997; Friesen et al., 2001). Clarion CII listeners generally performed similarly with 8-, 12-, and 16-channel experimental speech processors and their 8-channel clinical processor. Again, this similarity is probably due to the similarity in processing strategy (CIS) and the lack of difference between 8- and 16 channels. Thus, it seems most likely that Nulceus-24 listeners' lower performance with the experimental processors was due to the difference in stimulation strategy between their clinical processor (SPEAK or ACE) and the experimental procesors (CIS).

Another interesting difference between the experimental processors and clinical processors for Nucleus-24 listeners is in the average stimulation pulse rate. All of the Nucleus-24 listeners had the SPEAK processing strategy in their original processor. The average stimulation rate of this strategy is from 180 to 300 ppse (Cochlear Corporation, 1995). Although not a significant difference, Nucleus listeners tended to performed better with the 250-ppse processor than with the 900-ppse processor.

Overall, no listener was able to achieve significantly better speech recognition with any of the experimental speech processors than they obtained with their clinical processors. While some of the experimental processors allowed equivalent performance, even in laboratory testing with no practice. It is possible that some of the experimental processors could allow improved speech recognition following some time period for them to accommodate to the new processor. However, since the present results show no improvement in speech recognition as a function of stimulation rate there is no reason to suppose that accommodation would favor high rates more than lower rates.

Temporal Coding of Speech Features

Although high stimulation rates did not produce an increase in overall speech recognition, it is possible that high stimulation rates might have improved reception of speech features that depend heavily on temporal cues, such as consonant voicing and manner. Fu and Shannon (2000) measured consonant recognition as a function of stimulation rate, from 150 ppse to 500 ppse in Nucleus-22 cochlear implant patients listening to an experimental CIS processor. They noted that with manner and voicing, scores increased from 50 to 150 ppse and showed no further increase from 150 to 500 ppse. Consistent with their findings, information received for voicing and manner in the present study generally remained stable for all stimulation rates, except for a decrease at the lowest stimulation rate of 200 ppse with the Clarion C1. Thus, it appears that high stimulation rates do not improve either overall speech recognition, or reception of specific speech features that are more dependent of temporal information. Note that implant listeners with N-24 and Clarion C1 devices only received about 40% of the information on voicing and about half of the manner information. In normal hearing listeners Shannon et al. (1995) found reception of almost 100% of the information on voicing and manner with CIS-like noise-band processors with 2, 3, and 4 channels. It is surprising that implant listeners received so much less of the information on speech features that are thought to be primarily conveyed by temporal cues, and are received at essentially 100% with only two spectral channels in NH listeners. One possible explanation is that these temporal cues are not properly represented in implants. But we know that temporal cues are highly robust to distortion in amplitude (Fu and Shannon, 1998), and that the stimulation rates used in the present experiment should have easily conveyed all temporal cues below 400 Hz, which should include all voicing and manner cues. Indeed, no significant improvement was observed in voicing and manner cues as stimulation rate was increased. To understand this difference we must look at what speech processor conditions interfere with reception of temporal cues. Shannon et al. (1998), using noise-band processors in NH listeners, saw a dramatic reduction in the reception of voicing and manner cues when the frequency-place mapping was distorted. In their experiment, voicing, manner and place information received were all reduced to single channel levels when the frequency-place mapping of a four-channel processor was warped by a logarithmic transformation. Thus, it is possible that distortion in the mapping of frequency information to cochlear place is reducing implant listeners' ability to make full use of temporal cues. We know that the frequency-place mapping in cochlear implants is affected by the electrode insertion depth, and there is an inherence frequency-place compression in the normal clinical mapping process. Additional distortion in the frequency-place map might occur due to current flow in the cochlea and uneven nerve survival.

CONCLUSIONS

Overall, little difference was observed in speech understanding performance as a function of stimulation rate in three cochlear implant devices: the Clarion C1, the Clarion CII, and the Nucleus-24. In the CLARION C1 device performance increased from 200 ppse to 400 ppse, but showed no further increase from 400 to 1600 ppse. For the Nucleus-24 and the Clarion CII devices, scores remained relatively stable for all stimulation rates, from 250 ppse to 15475 ppse. In all devices speech recognition was significantly poorer for 4-electrode processors, than for 8-electrode and higher processors, but there was no significant difference in performance between 8-, 12-, and 16-electrode processors. These results are consistent with previous studies showing a lack of effect of stimulation rate on speech recognition and in the asymptote in performance for more than 8 electrodes.

REFERENCES

Brill, S., Gstöttner, W., Helms, J., v.Ilberg, C., Baumgartener, W., Müller, J., and Keifer, J. (1997). Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+, *American Journal of Otology Supplement*, 18, S104-S106.

Brill, S.M., Hochmair, I., and Hochmair, E.S. (1998). "The importance of stimulation rate in pulsatile stimulation strategies in cochlear implants," Presented at the *XXIV International Congress of Audiology*, Buenos Aires.

Burns, E.M. and Viemeister, N.F. (1976) Nonspectral Pitch, *J. Acoust. Soc. Amer.*, 60, 863-869.

Burns, E.M. and Viemeister, N.F. (1981) Played-again SAM: Further observations on the pitch of amplitude-modulated noise, *J. Acoust. Soc. Amer.*, 70, 1655-1660.

Chatterjee, M. and Robert, M.E. (2001). Noise enhances modulation sensitivity in cochlear implant listeners: Stochastic resonance in a prosthetic sensory system?, *Journal of the Association for Research in Otolaryngology*, 2(2)159-171.

Clarion by Advanced Bionics. *SCLIN 98 for windows device fitting manual.* Sylmar, California. 1998.

Cochlear Corporation (1995), Audiologist Handbook, Englewood, Colorado.

Drullman, R., Festen, J.M., and Plomp, R. (1994). “Effect of temporal envelope smearing on speech reception,” *J. Acoust. Soc. Am.* 95, 1053-1064.

Eddington, D.K., Dobelle, W.H., Brackmann, D.E., Mladevosky, M.G., and Parkin, J.L. (1978). Auditory prosthesis research with multiple channel intracochlear stimulation in man, *Ann. Otol. Rhinol. Laryngol.*, 87, Suppl. 53, 1-39.

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

Friesen, L., Shannon, R.V., Baskent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants, *Journal of the Acoustical Society of America*, 110(2), 1150-1163.

Fu, Q.-J. and Shannon, R.V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners, *Journal of the Acoustical Society of America*, 104(5), 2570-2577.

Fu, Q.-J. and Shannon, R.V. (2000). Effect of stimulation rate on phoneme recognition in cochlear implants, *Journal of the Acoustical Society of America*, 107(1), 589-597.

Fu, Q.-J., Galvin, J., and Wang, X. (2001). Time-altered sentence recognition by normal-hearing listeners and Nucleus-22 cochlear implant listeners, *J. Acoust. Soc. Amer*, in press.

Fu, Q-J. and Galvin, J. (2001). Spectrally asynchronous speech recognition by normal-hearing listeners and Nucleus-22 cochlear implant users, *J. Acoust. Soc. Amer.*, submitted.

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

Kiefer, J., von Illberg, C., Rupprecht, V., Huber-Egener, J., and Knecht, R. (2000). Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: Effect of variations in stimulation rate and number of channels, *Ann. Otol. Rhinol. Laryngol.*, 109, 1009-1020.

Lawson D.T., Wilson B.S., Zerbi M. & Finley C.C. (1996). "Speech processors for auditory prostheses", Third Quarterly Progress Report, NIH Contract N01-DC-5-2103.

Loizou, P.C., Poroy, O., and Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding, *J. Acoust. Soc. Am.*, 108(2), 790-802.

Miller, G. & Nicely, P. (1955). An analysis of perceptual confusions among some english consonants. *Journal of Acoustical Society of America*, 27, 338-352.

Nucleus Technical Reference Manual. Cochlear Corporation. Englewood, Colorado, 1995.

Pelizzone, M., Tinembart, J., Francois, J., Steger, D., Boex, C., and Montandon, P. (1998). "Very-high stimulation rates for cochlear implants," presented at the *4th European Symposium on Paediatric Cochlear Implantation*, s-Hertogenbosch, The Netherlands.

Plomp, R. (1983). The role of modulation in hearing, in Hearing – Physiological Bases and Psychophysics, R. Klinke and R. Hartmann, (Eds.), pp. 270-276, Springer-Verlag, Berlin.

Robert, ME. (1997). *AIPSS-ID: Phoneme Identification Software*. Los Angeles; House Ear Institute.

Rosen, S., Walliker, J., Brimacombe, J.A., and Edgerton, B.J. (1989). Prosodic and segmental aspects of speech perception with the House/3M single channel implant, *J. Speech Hear. Res.*, 32, 93-111.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory, and linguistic aspects, *Philos. Trans. Royal Soc. London Ser. B Biol. Sci.*, 336, 367-373

Rubinstein, J.T., Wilson, B.S., Finley, C.C., and Abbas, P.J. (1999). Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation, *Hearing Research*, 127, 108-118.

Saberi, K. and Perrott, D.R. (1999). Cognitive restoration of reversed speech, *Nature*, 398, 760.

Shannon, R.V. (1983) Multichannel electrical stimulation of the auditory nerve in man: I. Basic psychophysics, *Hearing Research*, 11, 157-189.

Shannon, R.V. (1992). Temporal modulation transfer functions in patients with cochlear implants, *J. Acoust. Soc. Amer.*, 91, 1974-1982

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

Shannon, R.V., Zeng, F.-G., and Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues, *J. Acoust. Soc. Amer.*, 104(4), 2467-2476.

Shannon, R.V., Jensvold, A., Padilla, M., Robert, M., and Wang, X. (1999). Consonant recordings for speech testing, *Journal of the Acoustical Society of America (ARLO)*, 106(6), L71-L74.

Tong, Y.C., Blamey, P.J., Dowell, R.C., and Clark, G.M. (1983) Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant, *J. Acoust. Soc. Amer.*, 74, 73-80.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R.L. (1987). "Pitch perception by cochlear implant subjects". *J. Acoust. Soc. Am.*, 82, 106-115.

Turner, C.W., Fabry, D.A., Barrett, S., and Horowitz, A.R. (1992). Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners, *J. Speech Hearing Res.*, 35, 942-949.

Vandali, A.E., Whitford, L.A., Plant, K.L., and Clark, G.M. (2000). "Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system", *Ear and Hearing*, 21(6), 608-624.

van der Horst, R., Leeuw, A.R., and Dreschler, W.A. (1999). Importance of temporal envelope cues in consonant recognition, *J. Acoust. Soc. Amer.*, 105(3), 1801-1809.

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

Van Tasell, D.J., Greenfield, D.G., Logemann, J.J., and Nelson, D.A. (1992). "Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants," *J. Acoust. Soc. Amer.* 92, 1247-1257.

Vischer M, Haenggeli A, Zhang J, Pelizzone M, Hausler R, Rouiller EM (1997), "Effect of high-frequency electrical stimulation of the auditory nerve in an animal model of cochlear implants ", *Am J Otol* 18(6 Suppl):S27-9.

Wilson. B. S., Finley, C. C, Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). New levels of speech recognition with cochlear implants, *Nature*, 352, 236-238.

Wilson, B.S., Finley, C.C., Lawson, D., and Zerbi, M. (1997a). Temporal representations with cochlear implants. *Am J Otol.* 18, S30-S34.

Wilson, B.S, Finley, C.C., Zerbi, M., Lawson, D.T., and van den Honert, C. (1997b). Speech processors for auditory prostheses, NIH project N01-DC-5-2103, Seventh Quarterly progress report, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD.

Wilson, B.S., Zerbi, M., Finley, C.C., Lawson, D., and van den Honert, C. (1997c). Speech processors for auditory prosthesis, NIH project N01-DC-5-2103, Eighth quarterly progress report, Neural Prosthesis Program, national Institutes of Health, Bethesda, MD.

Table 1: CI subject information.

Listener	Speech Processing Strategy	Age (Years)	Gender	CI Ear	Etiology	Age of HL Onset		Age of Profound HL Onset		Hearing Aid Usage		Duration of CI Use
						L	R	L	R	L	R	(Years)
C1	CIS	66	F	L	Otosclerosis	32	32	45	45	Y	N	1
C3	CIS	56	M	R	Unknown	18	0	18	45	N	N	3
C5	CIS	38	M	L	Unknown	3	3	28	22	Y	Y	2.5
C10	CIS	54	F	R	Hereditary	33	33	48	40	N	N	1
C11	CIS	50	F	R	Unknown	33	45	47	52	N	Y	1
C14	CIS	55	F	R	Hereditary Otosclerosis	34	34	52	52	Y	Y	1
C15	SAS	42	F	L	Meningitis	3	3	12	12	Y	Y	1
C16	CIS	49	M	R	Unknown	40	40	40	40	Y	Y	1
N24 1	SPEAK	66	F	R	Unknown	33	41	57	57	Y	Y	3
N24 2	SPEAK	50	F	R	Unknown	1	1	28	28	N	N	2
N24 3	SPEAK	58	M	R	Unknown	17	17	57	57	Y	Y	3
N24 4	SPEAK	63	F	L	Unknown	37	37	55	55	Y	Y	1

Table 2: Percent information received for consonants, averaged across stimulation rates > 400 ppse.

# electrodes	Device	Voicing	Manner	Place
16	N24	37.2	52.2	38.4
	CII	60.3	68.3	49.2
12	N24	37.9	52.1	37.5
	CII	67.3	70.3	46.7
8	N24	39.1	53.2	36
	C1	39.3	55.8	42.2
	CII	68.2	69.0	45.3
4	N24	35.5	50.8	29.3
	C1	38.2	55.6	40.8
	CII	62.9	60.9	40.7

Publications and Presentations in this Quarter

Peer-Reviewed Publications:

Galvin, J.G., and Baskent, D. (2001). Holes in hearing, J. Assoc. Res. Otolaryngol., 3(2), published online in December.

Hitselberger, W.E., Brackmann, D.E., Day, J.D., Shannon, R.V., Otto, S.A. and Ghosh, S. (2001) Auditory brainstem implants, in Operative Techniques in Neurosurgery, 4(1), 47-52.

Non-peer-reviewed Publications:

Friesen, L., Fu, Q.-J., Chatterjee, M., and Galvin, J.J.III (2001). "Cochlear implant research: Overview, current and future trends", ASHA Newsletter, Fall.

Shannon, R.V. (2001). Speech Perception, in McGraw-Hill 2002 Yearbook of Science and Technology, McGraw-Hill, New York, pp 333-336.

Manuscripts Submitted this Quarter:

Baskent, D. and Shannon, R.V. Speech recognition under conditions of frequency-place compression and expansion, J. Acoust. Soc. Amer., submitted Dec 01.

Chatterjee, M. Modulation masking in cochlear implant listeners: Envelope vs. tonotopic components, J. Acoust. Soc. Amer., submitted Dec 13.

Manuscripts Submitted in Previous Quarters and Manuscripts in Press:

Baskent, D. and Shannon, R.V. (2002). Speech recognition under conditions of frequency-place compression and expansion, J. Acoust. Soc. Amer., submitted Dec 01.

Chatterjee, M. (2002). Modulation masking in cochlear implant listeners: Envlope vs. tonotopic components, J. Acoust. Soc. Amer., submitted Dec 13.

Fu, Q.-J. and Shannon, R.V. Frequency mapping in cochlear implants, Ear and Hearing, in press.
Fu, Q.-J., Shannon, R.V., and Galvin, J. (2002). Performance over time of Nucleus-22 cochlear implant listeners wearing speech processor with a shifted frequency-to-electrode assignment, J. Acoust. Soc. Amer., (revision submitted Oct 01).
Rauschecker, J. and Shannon, R.V. (2002). Sending sound to the brain, Science, 295:1025-1029. (Invited Viewpoint)
Shannon, R.V. et al. (2002). The Auditory Brainstem Implant, in MIT Encyclopedia of Communication Disorders, R.D. Kent (Ed.), MIT Press, Cambridge, MA. (Invited chapter)
Zeng, F.G., Grant, G., Niparko, J., Galvin, J., Shannon, R.V., Opie, J., and Segel, P. (2002). Speech dynamic range and its effect on cochlear implant performance, J. Acoust. Soc. Amer., 111: 377-386.

Invited Presentations:

Shannon, R.V. (2001). Cochlear Implants and the Auditory System: Is New Technology Making a Difference?, Next Generation Cochlear Implants: Symposium and Practicum, Dept. of Otolaryngology, New York University School of Medicine, New York City, November 30-December 1. (Invited keynote speaker)

Presentations:

Shannon RV and Baskent D (2001). Speech Processors for Auditory Prostheses, Annual Progress Report, NIH Neural Prosthesis Workshop, Bethesda, MD, Oct 17-19. (oral presentation)
Baskent D and Shannon RV (2001). Speech recognition under conditions of frequency-place compression and expansion, NIH Neural Prosthesis Workshop, Bethesda, MD, Oct 17-19. (poster)
Padilla M and Shannon RV (2001). Speech recognition as a function of reduced spectral resolution and English experience in quiet and in noise, NIH Neural Prosthesis Workshop, Bethesda, MD, Oct 17-19. (poster)