

Effect of Stimulation Rate on Cochlear Implant Users' Phoneme, Word and Sentence Recognition in Quiet and in Noise

Robert V. Shannon Rachel J. Cruz John J. Galvin, 3rd

House Ear Institute, Los Angeles, Calif., USA

Key Words

Cochlear implant · Stimulation rate · Speech recognition · Speech processing

Abstract

High stimulation rates in cochlear implants (CI) offer better temporal sampling, can induce stochastic-like firing of auditory neurons and can increase the electric dynamic range, all of which could improve CI speech performance. While commercial CI have employed increasingly high stimulation rates, no clear or consistent advantage has been shown for high rates. In this study, speech recognition was acutely measured with experimental processors in 7 CI subjects (Clarion CII users). The stimulation rate varied between (approx.) 600 and 4800 pulses per second per electrode (ppse) and the number of active electrodes varied between 4 and 16. Vowel, consonant, consonant-nucleus-consonant word and IEEE sentence recognition was acutely measured in quiet and in steady noise (+10 dB signal-to-noise ratio). Subjective quality ratings were obtained for each of the experimental processors in quiet and in noise. Except for a small difference for vowel recognition in quiet, there were no significant differences in performance among the experimental stimulation rates for any of the speech measures. There was also a small but significant increase in subjective quality rating as stimulation rates increased from 1200 to 2400 ppse in noise. Consistent with previous studies, performance significantly

improved as the number of electrodes was increased from 4 to 8, but no significant difference showed between 8, 12 and 16 electrodes. Altogether, there was little-to-no advantage of high stimulation rates in quiet or in noise, at least for the present speech tests and conditions.

Copyright © 2010 S. Karger AG, Basel

Introduction

Cochlear implant (CI) users can achieve high levels of speech understanding despite the limited acoustic cues preserved by CI signal processing. Numerous studies [Dorman et al., 1997, 1998; Fishman et al., 1997; Friesen et al., 2001, 2005; Fu et al., 1998; Hill et al., 1968; Loizou et al., 2000; Shannon et al., 1995; Spahr and Dorman, 2006] have demonstrated good speech understanding in quiet and even at moderate levels of noise, given the temporal envelope of the speech signal and a small number (4–8) of spectral channels. CI technology and signal processing have improved over the past decades, incorporating larger numbers of electrodes, different stimulation modes and a variety of processing strategies to extract and convey the most important acoustic features of speech. However, the limiting factors in CI performance remain the poor spectral resolution (whether due to the limited number of electrodes or channel interactions) and limited access to temporal cues. To improve the tem-

poral resolution, many commercial CI devices employ high stimulation rates. While low-to-moderate rates may provide adequate temporal cues for speech recognition under quiet, optimal listening conditions, high rates have been proposed [Rubinstein et al., 1999; Wilson et al., 1998, 2000] to offer advantages for difficult listening conditions such as speech in noise or competing speech, music appreciation, etc.

High rates have been shown to expand the dynamic range (DR) of electrical stimulation relative to low rates, primarily by reducing threshold stimulation levels [Hong and Rubinstein, 2003, 2006]. However, the intensity resolution has been shown to be similar between low and high rates [Kreft et al., 2004] despite the wider DR with high rates. High rates also offer better temporal sampling of the input acoustic signal. However, CI listeners may not utilize high-frequency envelope cues due to temporal processing limits. Normal-hearing and CI users' temporal modulation sensitivity is limited to approximately 300 Hz. Several previous studies [Fu and Shannon, 2000; Xu and Pfungst, 2003; Xu and Zheng, 2007; Xu et al., 2005] have shown little improvement in phoneme recognition with temporal envelope cues beyond approximately 20 Hz.

High-rate 'conditioning' pulse trains presented at sub-threshold levels have been proposed [Hong and Rubinstein, 2003, 2006; Rubinstein et al., 1999] to improve the representation of temporal fine structure and expand the electrical DR. Similarly, Litvak et al. [2003a-c] demonstrated that high-rate carriers can result in stochastic activity in single auditory nerve fibers. This 'pseudospontaneous' activity is found in normal hearing and may be desirable in electric stimulation as it may desynchronize phase locking across neural populations, potentially improving signal detection and channel independence. Electrically stimulated auditory nerve fibers exhibit abnormally strong phase locking to the temporal fine structure of the electrical signal [van den Honert and Stypulkowski, 1984]. This phase locking may interfere with the auditory system's ability to utilize the temporal fine structure. Wilson et al. [1997] demonstrated that the use of a high-rate conditioning stimulus produced intracochlear evoked potentials that better approximated normal neural responses to auditory stimulation. Taken together, these studies suggest that high stimulation rates in electric hearing can produce temporal neural response patterns that are more similar to those of a normally hearing ear.

In terms of speech recognition, high stimulation rates have produced mixed results. Kiefer et al. [2000] tested 13

Med-El users' recognition of monosyllabic words and 2-digit numbers with relatively high (1515 or 1730 pulses per second per electrode, or ppse) or low stimulation rates (600 ppse) and found significantly better consonant recognition with the higher rates. Brill et al. [1997] measured speech recognition (consonants, vowels, numbers, sentences) in 3 Med-El Combi listeners using stimulation rates ranging from 1515 to 9090 ppse. While high rates sometimes improved performance for some subjects, the authors found no clear or consistent rate effect. Loizou et al. [2000] found that Ineraid CI users' monosyllabic word and phoneme recognition improved as the stimulation rate was increased from 400 to 2100 ppse. Nie et al. [2006] found a significant improvement in consonant recognition in quiet for Med-El users when the stimulation rate for experimental 4-channel processors was increased from 1000 to 4000 ppse.

In contrast, several studies have shown little-to-no difference in speech understanding between low and high rates. Verschuur [2005] found no significant changes in CI users' speech performance (phonetic categorization, identification of phonemes, words and sentences) as a function of stimulation rate, although some subjects' performance improved with high rates. Similarly, Plant et al. [2007] found a preference for high stimulation rates in some subjects, but no significant difference in mean CI performance between low and high rates.

Lawson et al. [1996] found no significant difference in CI users' consonant recognition across 3 experimental rates (250, 833 and 2525 ppse). Several studies [Fu and Shannon, 2000; Kiefer et al., 2000; Wilson et al., 2000] measured phoneme recognition as a function of stimulation rate and of the temporal envelope filter cutoff frequency, and found no significant effect of stimulation rate. Vandali et al. [2000] found poorer performance with high rates for monosyllable word recognition in quiet and sentence recognition in noise, largely driven by the performance of 1 of the 5 CI subjects. Friesen et al. [2005] found no significant difference in speech performance in quiet between experimental high-rate processors and CI subjects' low-rate clinical processors.

The range of outcomes across these previous studies may be due to differences in test materials (phonemes, words or sentences) or test conditions (quiet or noise), or differences across experimental processors. Interactions between the number of electrodes and the stimulation rate may have contributed to the variability in outcomes, as well as the effect of CI subjects' short- and long-term experience with the experimental high-rate processors.

Table 1. Subject demographics

Subject	Subject demographics						Baseline performance with clinical processors							
	age years	gen- der	etiology	CI ear	clinical strategy; rate ppse	CI use months	vowels		consonants		CNC		IEEE	
							quiet	10 dB SNR	quiet	10 dB SNR	quiet	10 dB SNR	quiet	10 dB SNR
CI-14	55	F	otosclerosis	R	CIS; 913.6	12	65	61	67	63	60	32	28	5
CI-15	42	F	meningitis	L	SAS; n/a	12	59	60	57	60	67	25	53	0
CI-16	49	M	sudden	R	CIS; 812.5	12	84	75	70	67	87	75	70	28
CI-17	53	M	sudden	L	MPS; 1625	15	58	47	61	63	68	53	48	18
CI-18	33	F	gradual	L	CIS; 928.6	15	81	74	73	68	85	61	68	38
CI-19	54	F	hereditary	R	CIS; 812.5	6	62	55	56	54	65	45	15	3
CI-20	59	F	gradual	L	SAS; n/a	8	55	32	55	42	33	18	10	0
Mean	49.3					11.4	66.3	57.7	62.7	59.6	66.4	44.1	41.7	13.1
SD	9.0					3.4	11.5	15.1	7.3	9.0	17.9	20.5	24.4	15.1

CNC = Consonant-nucleus-consonant; CIS = continuous interleaved sampling; SAS = simultaneous analog stimulation; MPS = multiple pulsatile stimulation.

In the present study, speech recognition performance was acutely measured in 7 CI subjects listening to 4-, 8-, 12- or 16-electrode speech processors, each mapped with target stimulation rates of 600, 1200, 2400 or 4800 ppse (except that the 4800-ppse rate could not be achieved with 16 electrodes). There are several differences in experimental conditions, processor mapping and subjects between the present study and the previous study by Friesen et al. [2005]. First, the present study measured phoneme, word and sentence recognition both in quiet and in steady speech-shaped noise at a +10 dB signal-to-noise ratio (SNR), whereas Friesen et al. [2005] measured performance only in quiet. Second, subjective quality ratings were obtained for each experimental processor in the present study but not in that by Friesen et al. [2005]. Third, the target stimulation rates were 600, 1200, 2400 or 4800 ppse. In Friesen et al. [2005], experimental rates ranged from 250 to 4901 ppse. The highest and lowest rates depended on the device and number of active electrodes, making it difficult to compare rate effects across the number of electrodes and across devices.

Methods

Subjects

All subjects were postlingually deafened CI users recruited from 3 implant centers in the Los Angeles area. The subject demographics are shown in table 1. All subjects gave their informed consent in accordance with local institutional review board requirements, and all subjects were paid for their participation.

Experimental Speech Processors

All subjects used the Advanced Bionics CII implant with the HiFocus electrode array and the Electrode Positioning System. The CII device contains 16 linearly spaced contacts and 16 independent current sources. The CII Bionic Ear™ Programming System (BEPS) [Advanced Bionics, 2001] was used to fit the 4-, 8-, 12- or 16-electrode experimental speech processors. All speech processors were fit with the continuous interleaved sampling (CIS) strategy [Wilson et al., 1993]. In all cases, monopolar stimulation was used and the pulse phase duration was fixed at 10.8 μs/phase. For each electrode map condition, experimental processors were created to achieve target stimulation rates of 600, 1200, 2400 and 4800 ppse by adjusting the interpulse interval from 0 to 321 μs. Because of interactions between the number of electrodes and the stimulation rate within the BEPS® software, the actual stimulation rate for the experimental processors varied somewhat from the target rates. Also, because of hardware and software limitations, only 4-, 8- and 12-electrode processors were tested for the 4800-ppse target rate. In total, 15 experimental processors were tested, along with subjects' clinical processors. The target and actual stimulation rates per electrode are shown in table 2.

Subjects were fit with experimental processors according to standard clinical fitting procedures. The gain (0 dB), clipping levels (2048 μA), volume range (lower limit = 100; upper limit = 0), input DR (50 dB) and radio frequency settings (0) were fixed at the default settings. For each processor, threshold (T) and most comfortable loudness (M) levels were obtained; the M levels were then loudness balanced across the active electrodes. Finally, the M levels were globally adjusted for loudness using continuous speech, i.e. HINT (Hearing in Noise Test) sentences [Nilsson et al., 1994], presented via a single loudspeaker at 70 dBA. For all experimental processors, the overall input acoustic frequency range was 350–5500 Hz. As the number of electrodes was reduced from 16 to 4, the acoustic input frequency range was divided into a smaller number of logarithmically spaced frequency bands. The frequency-to-electrode assignments for the different processors are shown in figure 1.

Table 2. Target and actual stimulation rates for experimental processors

Target rate, ppse	Electrodes, n	Actual rate, ppse
600	4	725
	8	611
	12	595
	16	644
	mean	644
	SD	58
1200	4	1221
	8	1160
	12	1289
	16	1160
	mean	1208
	SD	61
2400	4	2900
	8	2900
	12	2578
	16	2900
	mean	2820
	SD	161
4800	4	5156
	8	5156
	12	3867
	16	
	mean	4726
	SD	744

Speech Materials and Testing Procedures

Speech performance was measured in quiet and in steady speech-shaped noise (+10 dB SNR) for medial vowel recognition, medial consonant recognition, consonant-nucleus-consonant (CNC) monosyllabic word recognition and IEEE sentence recognition. All speech materials were presented in a soundfield at 70 dBA via a single loudspeaker (Grason Stadler). The subjects were tested while seated in a double-walled, sound-treated room (IAC), directly facing the loudspeaker (1 m distance). The subjects were familiarized with the test procedures and software using their clinical processors. The experimental processors were tested immediately after fitting with no practice or familiarization. All tests were first conducted in quiet and then in steady speech-shaped noise (+10 dB SNR). The test order of experimental processors was randomized within and across subjects, and the speech test order was randomized across experimental processors and across subjects.

The vowel materials consisted of 12 tokens produced by 10 talkers (5 male and 5 female) presented in a /hVd/ context (/i, ɪ, e, ε, æ, a, ɔ, o, u, ʌ, ɜ, ɝ/ or 'heed', 'hid', 'hayed', 'head', 'had', 'hod', 'hawed', 'hoed', 'hood', 'who'd', 'hud', 'heard'). Vowel tokens were

selected from the stimuli recorded by Hillenbrand et al. [1995]. Vowel recognition was measured using a closed-set, 12-alternative forced-choice procedure. Custom software [Robert, 1997] was used to deliver the stimuli, collect subject responses and calculate the information received [Miller and Nicely, 1955]. During testing, a vowel token was randomly selected (without replacement) from the stimulus set. The subjects were asked to click on the response that matched the stimulus; 12 response buttons were shown onscreen, labeled in a /hVd/ context.

The consonant materials consisted of 20 tokens produced by 10 talkers (5 male and 5 female) presented in a /aCa/ context (/b, d, g, p, t, k, m, n, l, r, f, v, s, z, ʃ, tʃ, ð, dʒ, w, j/ or 'aba', 'ada', 'aga', 'apa', 'ata', 'aka', 'ama', 'ana', 'ala', 'ara', 'afa', 'ava', 'asa', 'aza', 'asha', 'acha', 'atha', 'aja', 'awa', 'aya'). Consonant tokens were selected from the stimuli recorded by Shannon et al. [1999]. Consonant recognition was measured using a closed-set, 20-alternative forced-choice procedure. Again, custom software [Robert, 1997] was used to deliver the stimuli, collect subject responses and calculate the information received. During testing, a consonant token was randomly selected (without replacement) from the stimulus set. The subjects were asked to click on the response that matched the stimulus; 20 response buttons were shown onscreen, labeled in a /aCa/ context.

CNC materials [Peterson and Lehiste, 1962] consisted of 10 lists (50 words per list) of monosyllabic words produced by a single male talker. One list was tested for each experimental processor. Because the number of experimental processors exceeded the number of CNC word lists, some lists were repeated. The test order of the initial set of 10 lists was randomized; once this set was completed, another set of 10 lists was randomly generated. During testing, a word was randomly selected from the test list (without replacement). Subjects reported their answers orally to the researcher; if a reply was unintelligible, the subjects were asked to spell their answer. Performance was scored in terms of the percent of correctly identified words.

The IEEE materials [1969] consisted of 72 lists of sentences (10 sentences per list) produced by a single male talker. Note that the IEEE sentences are somewhat difficult, and generally much more difficult than HINT [Nilsson et al., 1994] sentences. Sentence stimuli were delivered and scored using custom software, i.e. HEISRT[®] [Fu and Shannon, 2000]. During testing, a sentence was randomly selected from the test list (without replacement). The subjects repeated the sentence as accurately as possible. Performance was scored in terms of the percent of words correctly identified. Two lists were tested for each experimental processor, and performance was averaged across the 2 lists.

After having completed the speech tests for each experimental processor, the subjects were asked to rate the sound quality of the processor. They were asked: 'if the sound quality of your everyday speech processor was a '5' on a scale of 1 to 10, how would you rate the comparable sound quality of this map?'

Results

The full battery of speech tests was administered with the subjects' clinical speech processors before and after testing with the experimental processors. The initial

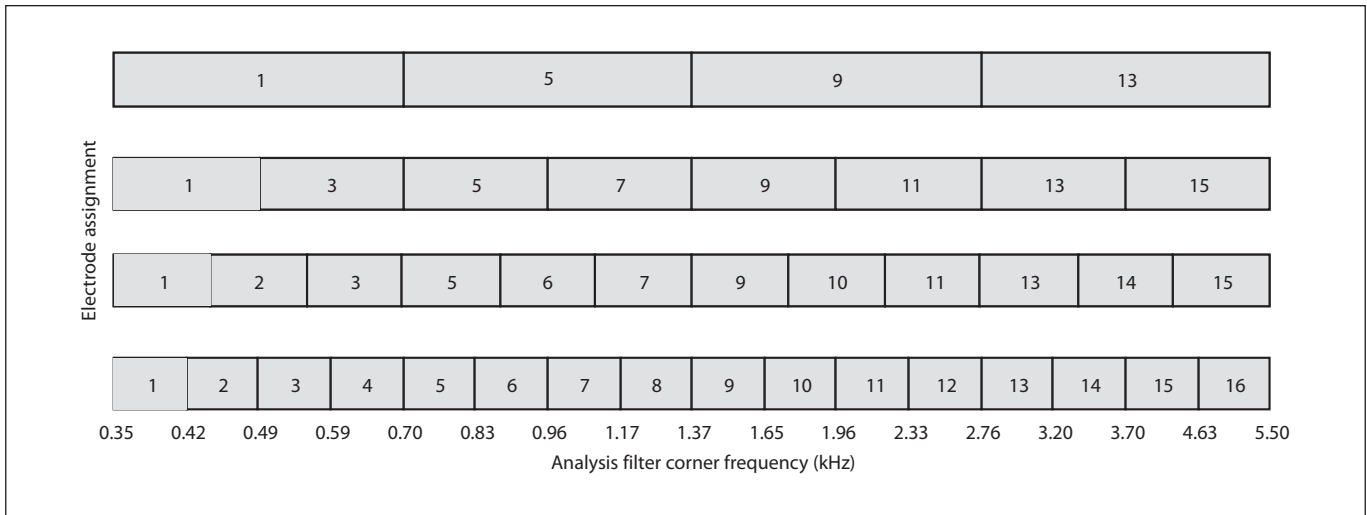


Fig. 1. Acoustic frequency-to-electrode assignment for experimental processors.

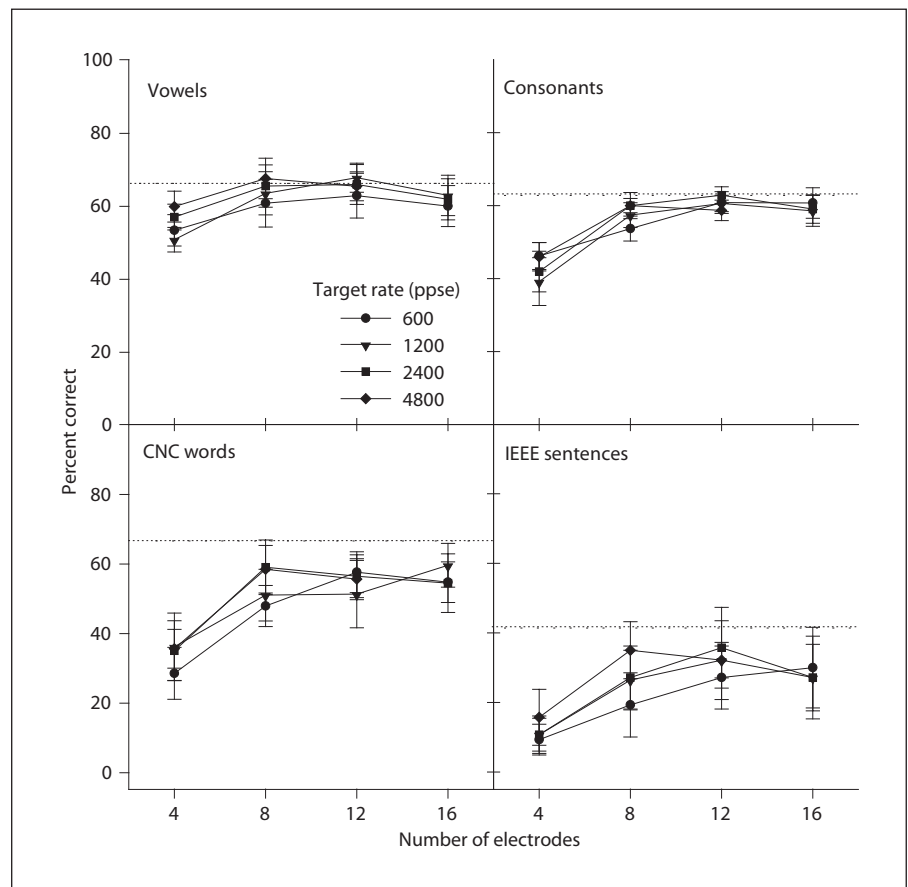


Fig. 2. Mean CI speech performance in quiet with experimental stimulation rates, as a function of the number of electrodes in the processor. Dotted line: mean performance with subjects' clinical processors. Error bars: ± 1 SEM.

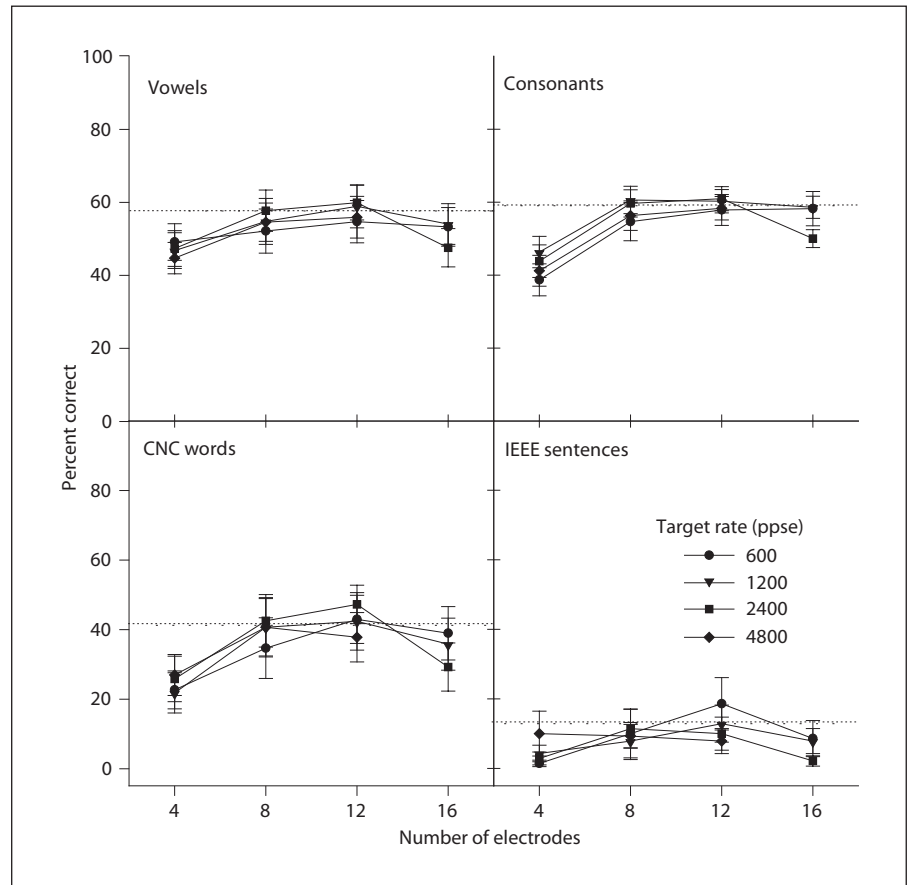


Fig. 3. Mean CI speech performance in steady, speech-shaped noise (+10 dB SNR) with experimental stimulation rates, as a function of the number of electrodes in the processor. Dotted line: mean performance with subjects' clinical processors. Error bars: ± 1 SEM.

Table 3. Results of two-way RM-ANOVA of speech performance with experimental processors, in quiet and in noise

Test	Electrode				Rate			
	d.f.; res	F ratio	p	post hoc (p < 0.05)	d.f.; res	F ratio	p	post hoc (p < 0.05)
<i>Quiet</i>								
Vowel	3; 56	4.84	0.012	12, 8>4	3; 56	6.01	0.005	4800>600
Consonants	3; 56	12.05	<0.001	16, 12, 8>4	3; 56	1.17	0.35	
CNC	3; 56	10.79	<0.001	16, 12, 8>4	3; 56	1.64	0.22	
IEEE	3; 56	9.12	<0.001	16, 12, 8>4	3; 56	2.65	0.08	
Quality	3; 56	14.97	<0.001	16, 12, 8>4	3; 56	1.89	0.17	
<i>Noise</i>								
Vowel	3; 56	5.37	0.008	12>4	3; 56	1.46	0.26	
Consonants	3; 56	29.60	<0.001	16, 12, 8>4	3; 56	2.55	0.09	
CNC	3; 56	11.92	<0.001	12, 8>4	3; 56	0.12	0.95	
IEEE	3; 56	6.69	0.003	12>16, 4	3; 56	0.85	0.48	
Quality	3; 56	14.36	<0.001	16>12, 8, 4, 12>4	3; 56	4.33	0.018	2400>1200

Post-hoc analyses were pair-wise Bonferroni comparisons. res = Residual.

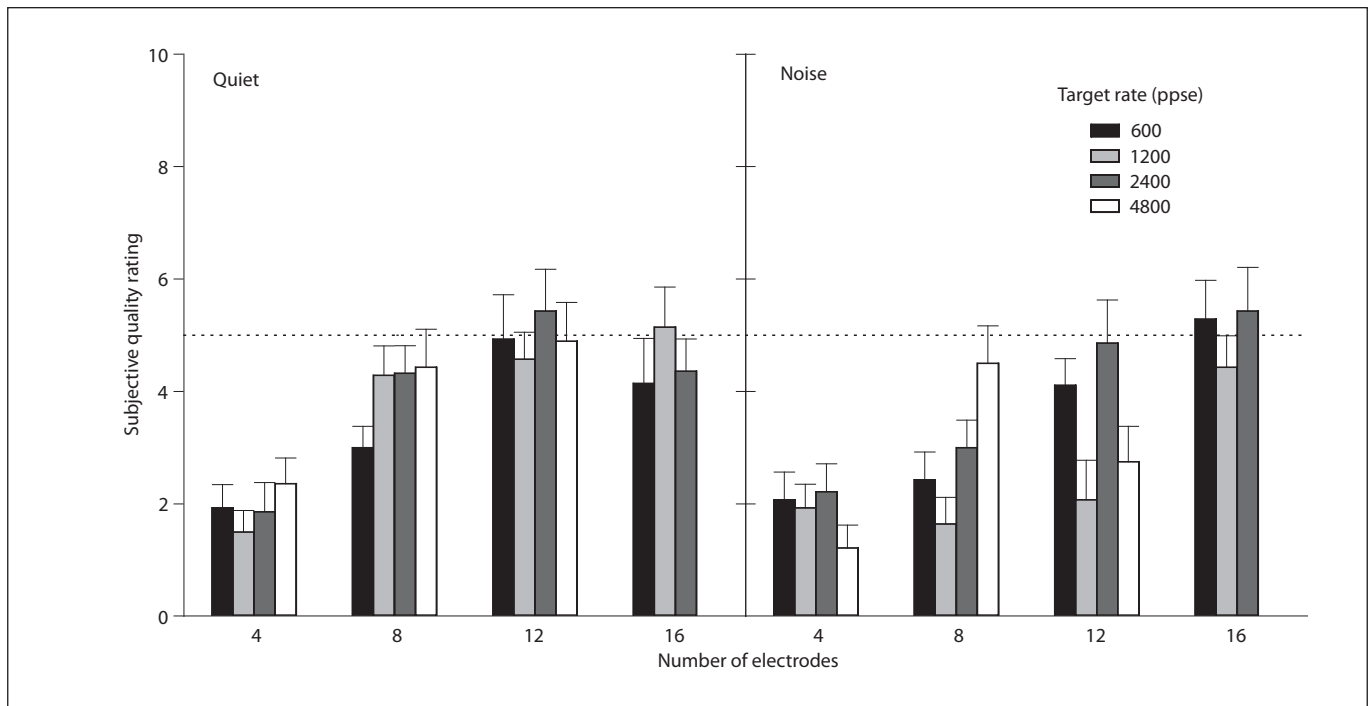


Fig. 4. Mean CI subjective quality ratings for experimental processors in quiet and noise, as a function of the number of electrodes in the processor. The subjects were asked to rate the quality of the processor on a scale of 1–10, with a ‘5’ rating (dotted line) corresponding to their clinical, everyday processor. Error bars: ± 1 SEM.

baseline measures served to familiarize the subjects with the test procedures, while the follow-up measures provided some indication of potential task-related learning effects. The mean performance (across baseline and follow-up measures) for each subject is shown in table 1. A two-way repeated measures analysis of variance (RM-ANOVA), with test date and speech test as factors and subject as the repeated measure, showed no significant difference between baseline and follow-up measures for any of the speech tests, in quiet [$F(1, 18) = 0.613$; $p = 0.463$] or in noise [$F(1, 18) = 0.849$; $p = 0.392$], suggesting there were no procedural learning effects during the course of the experiment.

Figure 2 shows the mean CI performance in quiet, as a function of the number of electrodes in the speech processor; the different symbols show the performance for the different stimulation rates. Table 3 shows the results of two-way RM-ANOVA (with number of electrodes and stimulation rate as factors and subject as the repeated measure) performed for each speech measure in quiet. There was a significant effect for the number of electrodes, primarily due to the poorer performance with 4

electrodes. There was no significant effect for stimulation rate, except for vowel recognition. Post-hoc Bonferroni comparisons found a small (5.23 percentage points) but significant difference only between the 600- and 4800-ppse rates.

Similarly, figure 3 shows the mean CI performance in steady noise (+10 dB SNR), as a function of the number of electrodes; the different symbols show the performance for the different stimulation rates. Table 3 also shows the results of the two-way RM-ANOVA (with number of electrodes and stimulation rate as factors and subject as the repeated measure) performed for each speech measure in noise. Similar to performance in quiet, there was a significant effect of the number of electrodes due to the poorer performance with 4 electrodes. There was no significant effect for stimulation rate.

Figure 4 shows subjective ratings for the experimental processors, as a function of the number of electrodes, in quiet and in noise. The results of the two-way RM-ANOVA are shown in table 3. There was a significant effect for the number of electrodes in quiet and in noise, largely due to the poorer rating with 4 electrodes. There

was no significant effect for stimulation rate in quiet. However, there was a significant effect for stimulation rate in noise due to the significantly higher rating for the 2400-ppse processor (rating: 3.85) relative to the 1200-ppse processor (rating: 2.52).

Discussion

Rate versus Electrode Effects

Previous studies have shown mixed results for high rates, possibly because the experimental rates were not adequately high. Wilson et al. [2000] suggested that rates beyond 4000 ppse might be necessary to induce the stochastic-like auditory nerve responses. In the present study, the experimental stimulation rates ranged from 596 to 5156 ppse. Except for the small difference (approx. 5 percentage points) between 600 and 4800 ppse for vowel recognition in quiet, we found no significant difference among the experimental rates in quiet or in noise for any of the speech measures or any of the spectral resolution conditions. The only significant rate effect was the subjective quality difference between the 1200-ppse and 2400-ppse processors in noise. Note that only CIS processors were tested, and that the +10 dB SNR may not have been challenging enough for some subjects, at least for the vowel and consonant tests. However, the +10 dB SNR was challenging enough to drop the average performance level obtained with the clinical processor from 66.4 to 44.1% for CNC words and from 41.4 to 12.9% for IEEE sentences.

Consistent with many previous studies [Dorman et al., 1998, 2000; Fishman et al., 1997; Friesen et al., 2001; Fu and Shannon, 1998; Loizou and Poroy, 2001; Loizou et al., 2000], there was a significant effect for the number of electrodes. Performance generally improved as the number of electrodes was increased from 4 to 8, beyond which there was no significant improvement. There was a slight drop in performance for some measures (e.g. vowels and IEEE sentences in quiet; vowels, CNC words and IEEE sentences in noise) as the number of electrodes was increased from 12 to 16, but this difference was not significant.

Given that listeners make greater use of temporal cues as the spectral resolution is reduced [Xu and Zheng, 2007; Xu et al., 2005], the 4-electrode processors would most likely be more sensitive to rate effects. Consonant recognition might also be more sensitive to rate effects than the more spectral cue-dependent or context-dependent vowel and sentence recognition measures. Also, speech un-

derstanding in noise might especially benefit from high rates as the additional temporal cues might help listeners segregate speech from noise. However, no significant effects of stimulation rate were observed for any of these conditions. Any of the physical advantages (temporal sampling) or physiological advantages (stochastic firing) associated with high rates did not provide any perceptual advantage, in quiet or in noise.

Subjective Evaluations

As they had only limited experience with the experimental processors, subjects consistently preferred the sound quality of their clinical processors to that of the experimental processors. In a similar study, Vandali et al. [2000] tested word and sentence recognition for 3 rates (250, 807 and 1615 ppse), then administered a subjective evaluation questionnaire after the subjects had continuously used the experimental processors over a period of several weeks. No clear trend in preference was found, even after several weeks of experience in a variety of listening conditions. While subjects preferred (on average) the 1615-ppse processor for music, they preferred the lower-rate processors for 6 other listening categories (including speech understanding in quiet and noise) as well as for the 'overall preference' category. Thus, in both a short-term adaptation study like that by Vandali et al. [2000] and in the present acute study, CI users did not explicitly prefer high-rate processors over clinical processors mapped with lower stimulation rates. However, longer-term studies [Plant et al., 2002, 2007; Vandali et al., 2000] have shown advantages of some stimulation rates (low, moderate or high) for individual CI patients, suggesting that CI users may need longer experience with a stimulation rate to develop a clear preference.

Why Do High Rates Not Improve Performance?

If high sampling rates provide some degree of physical and physiological advantage, why do they not provide a perceptual advantage? It is possible that the perceptual cues provided by these higher stimulation rates are subtle and can be utilized only with extensive experience. However, even the few studies that provided some adaptation period did not find any consistent advantage with high rates. It may be that the additional information provided by high rates is not perceptually salient. Perceptual rate discrimination becomes poor above 300 Hz for most listeners. Previous studies [Fu and Shannon, 2000; Xu and Pfungst, 2003; Xu and Zheng, 2007; Xu et al., 2005] also showed little decrement in performance when temporal information above 20 Hz is removed. While high rates

might offer improved temporal sampling, this extra information may not improve speech understanding beyond that already achieved with a much lower rate.

The detection of amplitude modulation in CI users has been shown to be better with low carrier rates, especially at low listening levels [Galvin and Fu, 2005, 2009; Pfungst et al., 2007]. The lower portion of the DR is an important perceptual region for low-amplitude consonant information, for which temporal cues might be important. Differences in loudness growth between low and high stimulation rates may explain some of the poorer modulation sensitivity with high rates at low levels [Galvin and Fu, 2009]. The electrode DR is larger with high rates, primarily due to lower detection thresholds. However, loudness grows more slowly with amplitude at high rates, especially in the lower portion of the DR [Rubinstein and Hong, 2003]. Indeed, differences in loudness growth between low and high rates may explain why the number of intensity just noticeable differences remains constant across rates [Kreft et al. 2004] despite the larger DR at high rates. Thus, greater changes in relative amplitude may be necessary for modulation detection with high-rate carriers. Galvin and Fu [2009] also found that high rates provided poorer modulation sensitivity at modulation frequencies ranging from 20 to 100 Hz, suggesting that CI subjects were unable to utilize the better temporal sampling with high rates. These single-channel modulation studies suggest that, if anything, speech performance might worsen with high rates.

With electrical stimulation, auditory nerve firing is strongly affected by the stimulus intensity and the refractory properties of the neurons. At rates below 800 ppse, neurons are abnormally highly phase-locked; neurons may fire at the same phase in every cycle of the stimulus. At higher stimulus rates (800–2000 ppse), there is a trade-off between stimulus intensity and neural refractory effects that weakens phase locking to the stimulus. Studies in cats and humans [Matsuoka et al., 1998; Wilson et al., 1997, 2000] demonstrated neuronal responses to high-rate electrical stimulation that were less deterministic and more like the natural acoustic stochastic response. High-rate, subthreshold conditioner pulse trains have been added to commercial CI devices to induce this stochastic response. While improved delivery of temporal information may be possible with high rates, the limited spectral resolution in CI devices may be the more limiting factor, relegating any improved temporal resolution with high rates to a negligible effect.

Increased electrode interaction with high rates might also offset potential advantages in temporal representa-

tions [Middlebrooks, 2004]. In rapid sequential stimulation of adjacent electrodes (as in the commonly used CIS processors), portions of a neural population may not fully recover between pulses due to refractory effects. Thus, temporal envelopes applied to adjacent electrodes may be spatially smeared, reducing the utility of temporal envelope information. With low rates, auditory neurons may more fully recover between successive pulses, and therefore may better preserve temporal envelope information. In the present study, there was no significant difference in performance across rates in the 4-electrode condition, which would presumably be least susceptible to electrode interactions. Only the lowest experimental rate (600 ppse) was beyond the neural refractory period; it is possible that rates below 600 ppse might have shown an effect. The sometimes poorer performance with 16 electrodes (relative to 12 or 8 electrodes) may have been due to increased electrode interaction; because there was no significant effect for stimulation rate, it is unclear whether rate-induced electrode interactions might have contributed to this small deficit in performance.

It is also possible that the present speech measures in quiet and in steady noise were not sensitive to the additional temporal cues which high rates might provide. High rates that provide better temporal sampling may be useful in providing better F0 cues, which are important when listening in competing speech or in dynamic noise, and for listening to music. More psychophysical research with multichannel stimulation may better show the limits of temporal cue perception, as well as effects of the stimulation rate on more complex perceptual situations. However, it is likely that other stimulation parameters (e.g. stimulation mode, electrode location, electrode design, frequency allocation, etc.) may more strongly influence performance, and CI research and development should perhaps more strongly commit to optimizing these parameters. The increase in overall stimulation rate allows for a greater number of electrodes to be stimulated per unit of time. This may be especially beneficial for current-focusing strategies (e.g. tripolar, quadripolar) that require multiple electrodes to be stimulated simultaneously or sequentially.

Summary and Conclusions

CI users' phoneme, word and sentence recognition in quiet and in steady noise (+10 dB SNR) was tested for a range of target stimulation rates (600–4800 ppse), and for a range of spectral resolution conditions (4–16 elec-

trodes). Except for vowel recognition in quiet, there were no significant differences among the experimental stimulation rates for any of the spectral conditions or for any of the speech tests in quiet or in noise. Consistent with previous studies, performance improved as the number of electrodes was increased from 4 to 8, beyond which there was no improvement. CI subjects generally rated the experimental processors as sounding poorer than their clinical processors, especially in noise. There were no significant differences in subjective quality ratings for the experimental stimulation rates, except between the 1200- and 2400-pps rates in quiet and in noise. Overall, the results suggest that the putative advantages associated with high rates (e.g. lower thresholds, expanded DR, bet-

ter temporal sampling, stochastic firing, etc.) did not benefit CI performance, at least for the present speech tests and conditions.

Acknowledgments

Thanks to our research subjects for their willing participation and their commitment to completing this study through many months of testing. Thanks also to Mark Robert for use of the Conдор[®] software, Dr. Qian-Jie Fu for use of the HEISRT software, and Advanced Bionics Corporation, especially Dr. Leo Litvak and Phil Siegel, for their technical assistance and for use of the BEPS software. This study was supported by NIH grant No. 5 R01 DC 01526.

References

- Advanced Bionics CII Bionic Ear Programming System. Valencia, Advanced Bionics Corporation, 2001.
- Brill SM, Gstöttner W, Helms J, von Ilberg C, Baumgartner W, Müller J, Kiefer J: Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+. *Am J Otol* 1997;18(suppl):S104–S106.
- Dorman MF, Loizou PC, Fitzke J: The identification of speech in noise by cochlear implant patients and normal-hearing listeners using 6-channel signal processors. *Ear Hear* 1998; 19:481–484.
- Dorman MF, Loizou PC, Fitzke J, Tu Z: Recognition of monosyllabic words by cochlear implant patients and by normal-hearing subjects listening to words processed through cochlear implant signal processing strategies. *Ann Otol Rhinol Laryngol Suppl* 2000; 185:64–66.
- Dorman MF, Loizou PC, Rainey D: 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* 1997;102:2403–2411.
- Fishman KE, Shannon RV, Slattery WH: Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res* 1997;40:1201–1215.
- Friesen LM, Shannon RV, Baskent D, Wang X: Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am* 2001;110:1150–1163.
- Friesen LM, Shannon RV, Cruz RJ: Effects of stimulation rate on speech recognition with cochlear implants. *Audiol Neurootol* 2005; 10:169–184.
- Fu QJ, Shannon RV: Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners. *J Acoust Soc Am* 2000; 107:589–597.
- Fu QJ, Shannon RV, Wang X: Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *J Acoust Soc Am* 1998;104:3586–3596.
- Galvin JJ 3rd, Fu QJ: Effects of stimulation rate, mode, and level on modulation detection by cochlear implant users. *J Assoc Res Otolaryngol* 2005;6:269–279.
- Galvin JJ 3rd, Fu QJ: Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users. *Hear Res* 2009;250:46–54.
- Hill FJ, McRae LP, McClellan RP: Speech recognition as a function of channel capacity in a discrete set of channels. *J Acoust Soc Am* 1968;44:13–18.
- Hillenbrand J, Getty LA, Clark MJ, Wheeler K: Acoustic characteristics of American English vowels. *J Acoust Soc Am* 1995;97(pt 1):3099–3111.
- Hong RS, Rubinstein JT: High-rate conditioning pulse trains in cochlear implants: dynamic range measures with sinusoidal stimuli. *J Acoust Soc Am* 2003;114(pt 1):3327–3342.
- Hong RS, Rubinstein JT: Conditioning pulse trains in cochlear implants: effects on loudness growth. *Otol Neurotol* 2006;27:50–56.
- IEEE: IEEE Recommended Practice for Speech Quality Measurements. New York, Institute of Electrical and Electronic Engineers, 1969.
- Kiefer J, von Ilberg C, Rupperecht V, Hubner-Egner J, Knecht R: 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* 2000;109:1009–1020.
- Kreft HA, Donaldson GS, Nelson DA: Effects of pulse rate and electrode array design on intensity discrimination in cochlear implant users. *J Acoust Soc Am* 2004;116:2258–2268.
- Lawson DT, Wilson BS, Zerbi M, Finley CC: Speech processors for auditory prostheses. Third Quarterly Progress Report. NIH Contract N01-DC-5-2103. Research Triangle Park, Center for Auditory Prosthesis Research, 1996.
- Litvak L, Delgutte B, Eddington D: Improved neural representation of vowels in electric stimulation using desynchronizing pulse trains. *J Acoust Soc Am* 2003a;114(pt 1): 2099–2111.
- Litvak LM, Delgutte B, Eddington DK: Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains. *J Acoust Soc Am* 2003b;114(pt 1):2079–2098.
- Litvak LM, Smith ZM, Delgutte B, Eddington DK: Desynchronization of electrically evoked auditory nerve activity by high-frequency pulse trains of long duration. *J Acoust Soc Am* 2003c;114(pt 1):2066–2078.
- Loizou PC, Poroy O: Minimum spectral contrast needed for vowel identification by normal hearing and cochlear implant listeners. *J Acoust Soc Am* 2001;110(pt 1):1619–1627.
- Loizou PC, Poroy O, Dorman M: The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am* 2000;108:790–802.
- Matsuoka AJ, Abbas PJ, Rubinstein JT, Miller CA: The neurophysiological effects of stimulated auditory prosthesis stimulation. Seventh Quarterly Progress Report. NIH Contract N01-DC-6-2111. Iowa City, University of Iowa, 1998.

- Middlebrooks JC: Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *J Acoust Soc Am* 2004;116:452–468.
- Miller GA, Nicely PE: An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 1955;27:338–352.
- Nilsson M, Soli SD, Sullivan JA: Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 1994;95:1085–1099.
- Nie KB, Barco A, Zeng FG: Spectral and temporal cues in cochlear implant speech processing. *Ear Hear* 2006;27:208–217.
- Peterson GE, Lehiste I: Revised CNC lists for auditory tests. *J Speech Hear Disord* 1962;27:62–70.
- Pfingst BE, Xu L, Thompson CS: Effects of carrier pulse rate and stimulation site on modulation detection by subjects with cochlear implants. *J Acoust Soc Am* 2007;121:2236–2246.
- Plant KL, Holden L, Skinner M, Arcaroli J, Whitford L, Law M, Nel E: Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system. *Ear Hear* 2007;28:381–393.
- Plant KL, Whitford LA, Psarros CE, Vandali AE: Parameter selection and programming recommendations for the ACE and CIS speech processing strategies in the Nucleus 24 cochlear implant system. *Cochlear Implants Int* 2002;3:104–125.
- Robert ME: Condor ID. Los Angeles, House Ear Institute, 1997.
- Rubinstein JT, Hong R: Signal coding in cochlear implants: exploiting stochastic effects of electrical stimulation. *Ann Otol Rhinol Laryngol* 2003;191:14–19.
- Rubinstein JT, Wilson BS, Finley C, Abbas PJ: Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res* 1999;127:108–118.
- Shannon RV, Jensvold A, Padilla M, Robert ME, Wang X: Consonant recordings for speech testing. *J Acoust Soc Am* 1999;106:L71–L74.
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M: Speech recognition with primarily temporal cues. *Science* 1995;270:303–304.
- Spahr AJ, Dorman MF: Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. *Arch Otolaryngol Head Neck Surg* 2006;130:624–628.
- Vandali AE, Whitford LA, Plant KL, Clark GM: Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear* 2000;21:608–624.
- van den Honert C, Stypulkowski PH: Physiological properties of the electrically stimulated auditory nerve. 2. Single fiber recordings. *Hear Res* 1984;14:225–243.
- Verschuur CA: Effect of stimulation rate on speech perception in adult users of the Med-El CIS speech processing strategy. *Int J Audiol* 2005;44:58–63.
- Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M: Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehabil Res Dev* 1993;30:110–116.
- Wilson BS, Finley CC, Lawson DT, Zerbi M: Temporal representations with cochlear implants. *Am J Otol* 1997;18(suppl):S30–S34.
- Wilson BS, Lawson DT, Zerbi M, Finley C, van den Honert C: Speech processors for auditory prosthesis. Eighth Quarterly Progress Report. NIH Contract N01-DC-5-2103. Research Triangle Park, Center for Auditory Prosthesis Research, 1998.
- Wilson BS, Wolford R, Lawson DT: Speech processors for auditory prosthesis. Seventh Quarterly Progress Report. NIH Contract N01-DC-8-2105. Research Triangle Park, Center for Auditory Prosthesis Research, 2000.
- Xu L, Pfingst BE: Relative importance of temporal envelope and fine structure in lexical-tone perception. *J Acoust Soc Am* 2003;114:3024–3027.
- Xu L, Thompson CS, Pfingst BE: Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am* 2005;117:3255–3267.
- Xu L, Zheng Y: Spectral and temporal cues for phoneme recognition in noise. *J Acoust Soc Am* 2007;122:1758–1764.