Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants

Lendra M. Friesen,[a] Robert V. Shannon, Deniz Baskent, and Xiaosong Wang
Department of Auditory Implants and Perception, House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

(Received 9 September 2000; accepted for publication 3 May 2001)

Speech recognition was measured as a function of spectral resolution (number of spectral channels) and speech-to-noise ratio in normal-hearing (NH) and cochlear-implant (CI) listeners. Vowel, consonant, word, and sentence recognition were measured in five normal-hearing listeners, ten listeners with the Nucleus-22 cochlear implant, and nine listeners with the Advanced Bionics Clarion cochlear implant. Recognition was measured as a function of the number of spectral channels (noise bands or electrodes) at signal-to-noise ratios of +15, +10, +5, 0 dB, and in quiet. Performance with three different speech processing strategies (SPEAK, CIS, and SAS) was similar across all conditions, and improved as the number of electrodes increased (up to seven or eight) for all conditions. For all noise levels, vowel and consonant recognition with the SPEAK speech processor did not improve with more than seven electrodes, while for normal-hearing listeners, performance continued to increase up to at least 20 channels. Speech recognition on more difficult speech materials (word and sentence recognition) showed a marginally significant increase in Nucleus-22 listeners from seven to ten electrodes. The average implant score on all processing strategies was poorer than scores of NH listeners with similar processing. However, the best CI scores were similar to the normal-hearing scores for that condition (up to seven channels). CI listeners with the highest performance level increased in performance as the number of electrodes increased up to seven, while CI listeners with low levels of speech recognition did not increase in performance as the number of electrodes was increased beyond four. These results quantify the effect of number of spectral channels on speech recognition in noise and demonstrate that most CI subjects are not able to fully utilize the spectral information provided by the number of electrodes used in their implant. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1381538]

PACS numbers: 43.71.Ky, 43.66.Sr, 43.66.Ts [CWT]

I. INTRODUCTION

Previous work with cochlear implants has demonstrated that speech recognition increases with increasing number of electrodes (Holmes *et al.*, 1987; Dorman *et al.*, 1989; Kileny *et al.*, 1992; Geier and Norton, 1992; Lawson *et al.*, 1993; Collins *et al.*, 1994; Lawson *et al.*, 1996; Fishman *et al.*, 1997; Eddington *et al.*, 1997). However, most of this work has been done in quiet listening conditions, while most everyday listening situations contain background noise, which reduces intelligibility even for individuals with normal hearing. In noisy listening conditions even normal-hearing (NH) listeners with 16 spectral channels do not achieve the same performance level as with full-spectrum speech (Fu *et al.*, 1998; Eddington *et al.*, 1997).

Studies with cochlear implants (Fishman *et al.*, 1997; Eddington *et al.*, 1997; Dorman and Loizou, 1997, 1998; Dorman *et al.*, 1997; Fu *et al.*, 1998) have demonstrated that speech recognition improves as the number of electrodes increases in quiet listening conditions, at least up to four to seven electrodes. Fu *et al.* (1998) measured recognition of vowels and consonants as a function of signal-to-noise ratio in three cochlear implant listeners and in four normal-hearing listeners in conditions simulating cochlear implants with both CIS and SPEAK-like strategies. Recognition scores for vowels and consonants decreased as the S/N level worsened in all conditions. Recognition of vowels and consonants was further measured in Nucleus-22 cochlear implant users with either their normal SPEAK speech processor or a custom processor with a four-channel CIS strategy. The best cochlear implant users showed similar performance with the CIS strategy in quiet and in noise to that of normal-hearing listeners when listening to correspondingly spectrally degraded speech, suggesting that the noise susceptibility of cochlear implant users is at least partly due to the loss of spectral resolution. Eddington *et al.* (1997) found that three implant listeners with the Ineraid device with six-channel CIS processors were recognizing consonants and sentences in both quiet and in noise at the same level as normal-hearing listeners with four to six channel noise processors.

Studies with acoustic hearing have demonstrated that speech recognition is reduced when the spectral resolution is degraded by spectral smearing or hearing impairment (Stelmachowicz *et al.*, 1985; Dubno and Dorman, 1987; Horst, 1987; ter Keurs *et al.*, 1992, 1993; Hill *et al.*, 1968; Baer and Moore, 1993, 1994; Turner *et al.*, 1999; Shannon *et al.*,

[a]Author to whom correspondence should be addressed. Electronic mail: lfriesen@hei.org

TABLE I. General information on Nucleus-22 listeners (CI=cochlear implant, HL=hearing loss, C. Otosclerosis=Cochlear otosclerosis).

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 (years)
						L	R	L	R	L	R	
N3	SPEAK	56	M	R	Trauma	45	10	45	45	N	N	7
N4	SPEAK	40	M	R	Trauma	35	35	35	35	N	N	5
N6	SPEAK	65	F	R	Ototoxicity	54	54	54	54	Y	Y	7
N7	SPEAK	55	M	R	Unknown	20	20	47	44	Y	N	2
N9	SPEAK	55	F	L	Hereditary	8	8	38	38	Y	Y	7
N14	SPEAK	63	M	R	Unknown	37	37	47	61	N	Y	1
N15	SPEAK	70	F	L	C. Otosclerosis	62	62	75	75	Y	Y	2
N17	SPEAK	71	F	R	Unknown	41	41	68	68	Y	Y	1
N18	SPEAK	77	F	R	Otosclerosis	40	40	45	45	Y	Y	1
N19	SPEAK	70	M	L	Unknown	40	40	62	56	Y	N	1

1995; Boothroyd *et al.*, 1996; Dorman *et al.*, 1997; Dorman and Loizou, 1997, 1998; Nejime and Moore, 1997; Eddington *et al.*, 1997; Fu *et al.*, 1998). In general, these studies found that speech recognition in quiet listening conditions was highly resistant to spectral smearing, with significant decreases in performance occurring only when the spectrum was smeared over 1000 Hz, or reduced to less than four spectral channels. Speech recognition was more susceptible to spectral smearing in the presence of added noise (Fu *et al.*, 1998; Baer and Moore, 1993; Nejime and Moore, 1997; Eddington *et al.*, 1997).

In the present experiment, speech recognition was measured as a function of the number of electrodes in various levels of noise, for three processing strategies: SPEAK, CIS, and SAS. Speech recognition was also measured in normal-hearing listeners with noise-band processors (Shannon *et al.*, 1995) as a function of the number of bands and signal-to-noise ratio.

II. METHODS

A. Listeners

Ten adults (18 years and older) utilizing the Nucleus-22 cochlear implant with the SPEAK speech processing strategy and nine adults using the Clarion cochlear implant device, each having at least six months CI experience, participated in this study. Five of the Clarion patients used the continuous interleaved sampler (CIS) processor (Wilson *et al.*, 1991) and four used the simultaneous analog stimulation (SAS) processor. All were postlingually deafened and native speakers of American English. General demographic information for the 19 subjects is presented in Tables I and II. All Nucleus-22 listeners had 20 active electrodes available for use, while Clarion users had either seven or eight, depending on the speech processing strategy used: SAS users had seven available electrode pairs and CIS users had eight available electrode pairs. Five normal-hearing listeners, ranging in age from 18 to 53 years, were recruited as controls.

B. Speech materials

Speech perception tests used 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, 1998). Ten presentations (five

TABLE II. General information of Clarion listeners (CI=cochlear implant, HL=hearing loss, C. Otosclerosis=cochlear otosclerosis).

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 (years)
						L	R	L	R	L	R	
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
C4	CIS	51	F	L	Meningitis	1.5	1.5	47	47	Y	N	2
C5	CIS	38	M	L	Unknown	3	3	28	22	Y	Y	2.5
C9	CIS	46	F	R	Ototoxicity	43	43	45	45	Y	Y	0.5
C2	SAS	72	M	R	C. Otosclerosis	30	30	69	69	Y	N	2
C6	SAS	61	F	R	Menieres	22	33	57	57	Y	Y	1
C7	SAS	82	M	R	Unknown	15	15	63	63	N	Y	2
C8	SAS	76	M	R	Unknown	18	64	75	64	Y	N	0.5

male and five female talkers) each of 12 medial vowels (/i ɔ ɛ u ɪ ʊ ʌ æ ɝ o ɑ e/) were presented in a /h/-vowel-/d/ context (heed, hawed, head, who'd, hid, hood, hud, had, heard, hoed, hod, hayed). Chance level on this test was 8.33% correct and the 95% confidence level was 13.4% correct.

Consonant stimuli were taken from materials created by Turner *et al.* (1992, 1999) and Fu *et al.* (1998). Consonant confusion matrices were compiled from 12 presentations (2 repetitions of 3 male and 3 female talkers) 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, 1998; Shannon *et al.*, 1999) and the confusion matrices were analyzed for information received on the production-based categories of voicing, manner, and place of articulation (Miller and Nicely, 1995). Chance performance level for this test was 7.14% correct, and the 95% confidence level was 11.1% correct.

The CNC Word Test from the Minimum Speech Test Battery for Adult Cochlear Implant Users CD (House Ear Institute and Cochlear Corporation, 1996) was used to evaluate open-set phoneme and word recognition. The CD contains ten lists of 50 monosyllabic words containing 150 phonemes. Listener responses were scored separately for words and phonemes correctly identified. Because there were more test conditions (25) than lists of words (10), the word lists used in the conditions with the poorest scores were repeated.

Recognition of words in sentences was measured using the Hearing in Noise Test (HINT) sentences (Nilsson *et al.*, 1994) 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 ten 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. Sentences were scored in terms of words correct.

C. Experimental speech processor conditions

Each listener was tested with five experimental speech processors immediately after receiving them (no practice). Each of the five experimental processors was tested in quiet and with four different signal-to-noise ratios (S/N) of +15, +10, +5, and 0 for a total of 25 conditions. The Nucleus-22 SPEAK processing strategy divides speech into 20 contiguous frequency bands and normally assigns the output of each band to one electrode pair. The listeners' original frequency band divisions were used. In the present experiment, processors were created with 2, 4, 7, 10, and 20 activated electrodes by assigning the output of more than one band to a single electrode pair. In the normal 20-electrode processor the output of analysis bands 1, 2, 3, 4, and 5 would normally be assigned to active electrodes 1, 2, 3, 4, and 5, respectively. In the present experiment a four-electrode experimental processor was created by assigning the outputs of all five bands to active electrode 3 only. In this case active electrodes 1, 2, 4, and 5 received no stimulation. When this assignment pattern was repeated along the entire electrode array the outputs of the 20 analysis filters were presented to only four active electrodes. In similar fashion, analysis filters were summed to create processors with ten, seven, four, and two active electrodes. In the seven-electrode condition the basal-most electrode pair was assigned only two frequency bands instead of three [see Fishman *et al.* (1997) for more details].

In the normal SPEAK processing strategy the acoustic signal is analyzed into 20 frequency bands and between six and ten frequency bands with the highest energy are selected for stimulation approximately every 4 ms (McDermott *et al.*, 1992a, b; Seligman and McDermott, 1995). The average pulse rate per electrode was higher in the experimental processors, because the activated electrodes received the output from more than one analysis filter band. For example, if an electrode pair was assigned to receive the output of three contiguous analysis bands (seven-electrode processor condition), then that electrode pair received a stimulation pulse if any of the three filter bands was selected for stimulation. If all three filter bands were selected for stimulation, the electrode pair would receive three pulses in that stimulus frame. Thus, as the number of electrodes was reduced, the effective stimulation pulse rate on each electrode pair was increased. ''Stimulus level'' coding was used, which changes the electrical stimulation level by changing both pulse amplitude and pulse phase duration (Cochlear Corp., 1995). At high stimulation levels the pulse duration is longer, which results in a lower overall pulse rate. All listeners were programmed in the bipolar-plus-one mode electrode pairing (BP+1) for both their normal processors and for all experimental conditions.

With the Clarion SAS and CIS speech processing strategies, the outputs of seven or eight frequency bands are normally directed to seven or eight electrode pairs assigned to those frequency bands (Clarion Reference Manual, 1998). With a reduction in the number of electrode pairs, the total frequency range remains the same, but the range for each electrode is broadened, with the exception of the two-electrode processor. With the two-channel processor only high- and low-frequency bands are transmitted, and the mid-frequency information is left out (Breeuwer and Plomp, 1984). Five electrode conditions were created where all seven or eight electrode pairs were utilized initially and then reduced to six, four, three, and two pairs. For the six-electrode condition electrodes 1, 2, 4, 5, 7, and 8 were used with the CIS processing strategy, while electrodes 1, 2, 3, 5, 6, and 7 were used with SAS processing. The four-electrode condition utilized electrodes 1, 3, 5, and 7; the three-electrode condition involved electrodes 1, 4, and 7; and the two-electrode condition involved electrodes 2 and 6. In the Clarion device the overall stimulation rate is held constant for the CIS processor. As the number of electrodes was reduced, the stimulation rate per electrode increased.

Normal-hearing listeners were tested using a noise band simulation of CIS-like processing (see Shannon *et al.*, 1995). Test conditions consisted of the five S/N ratios used for implant listeners, with up to four additional noise conditions: −2.5, −5, −7.5, and −10 dB S/N. Acoustic processors were designed with 20, 16, 12, 8, 6, 4, and 2 bands. For two, four, six, and eight band processors the same frequency divisions were used as the Clarion processor. For noise-band processors with more than eight bands, the entire frequency range from 100 Hz to 6 kHz was divided into equal parts in terms of cochlear distance in mm, using the cochlear tonotopic formula of Greenwood (1990). The envelope was extracted from each band by half-wave rectification and low-pass filtering at 160 Hz. This envelope signal was then used to modulate a wideband white noise, which was then bandpass filtered with the same filter set as was used on the original speech signal. The modulated noise bands were then summed and presented through a calibrated loudspeaker in a sound treated room (IAC). The speech-shaped masking noise was added to the speech signals at the desired speech-to-noise ratio prior to processing.

D. Procedure

During all testing the listener was seated 1 m 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 1-in microphone (Model #4144) at the location of the listener's head. All speech materials were recorded. A computer with a sound card (Turtle Beach Fiji), CD player, and a GSI audiometer (Model 16) were 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 for that test.

Threshold (T) and comfort (C) or most comfortable (M) loudness levels were measured separately for each experimental condition. The five experimental processors were presented to each listener in random order. Within each of the five experimental processor conditions, the four noise conditions were presented in random order, following the condition in quiet. The condition in quiet was always presented first in order to further familiarize the listener with the task. The battery of speech tests was administered to each listener immediately after they were given the experimental processor (no practice). The listener's normal settings were restored to the speech processor after each testing session until the listener returned for the next experimental condition, typically one week later. After the T and C level adjustments, Nucleus-22 listeners' were told to set their sensitivity level to the most comfortable position with the function switch set to normal (N). This setting was used during all the test conditions for that particular processor.

For the Nucleus-22 device electrical thresholds (T) and maximum acceptable loudness (C) levels were obtained using the Nucleus diagnostic and programming system with a personal computer and a dual processor interface with Cochlear Corporation 6.100 software. For obtaining T and C levels the stimulus was a 250-Hz pulse train of 500-ms duration. Threshold levels were estimated by a standard clinical bracketing procedure. 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 judged the loudness was at the maximum acceptable level. Adjacent electrodes were balanced for loudness at C level for each electrode.

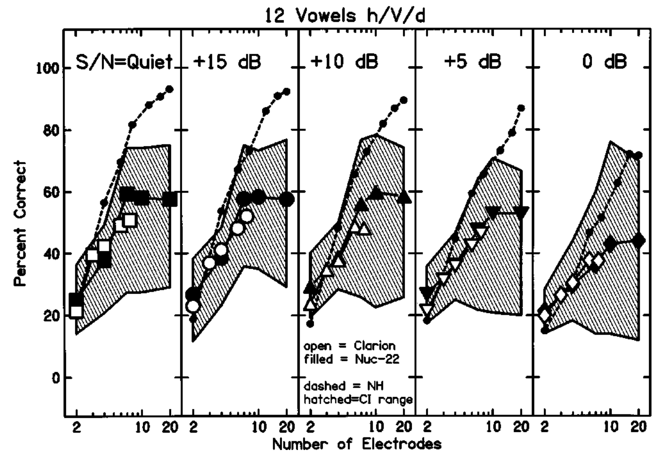

FIG. 1. Recognition of 12 medial vowels in a h/V/d context as a function of the number of spectral channels for normal-hearing listeners (dashed line with small filled symbols) or as a function of the number of electrodes used with Nucleus-22 cochlear implant listeners (filled symbols) and Clarion cochlear implant listeners (open symbols). The hatched area plots the range of performance across all 19 cochlear implant listeners. From left to right the panels present vowel recognition as a function of decreasing signal-to-noise ratio.

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 as in 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 judged 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.

The SAS measurement procedures were identical to CIS except the measurement of T and M levels began with the most basal channel, whereas with CIS the measurements began with the most apical channel as per the Clarion device fitting manual (Clarion, 1998).

III. RESULTS

Figures 1–4 present the results for vowels, consonants, CNC words, and HINT sentences, respectively. Within each figure, the panels present recognition performance for quiet listening conditions and signal-to-noise levels of +15, +10, +5, and 0 dB, respectively, from left to right. In each panel

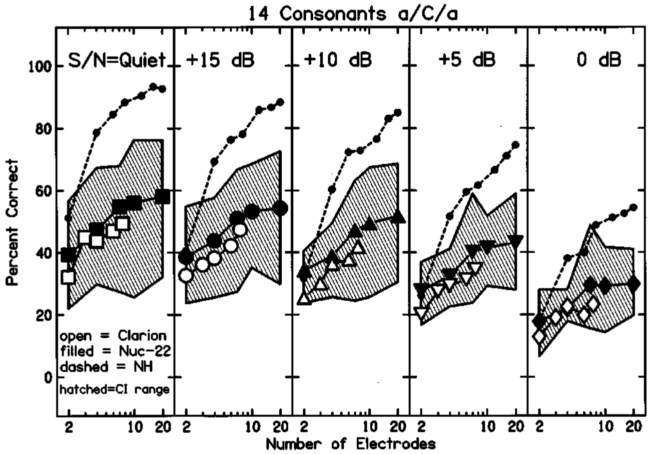

FIG. 2. Recognition of 14 medial consonants in an a/C/a context as a function of the number of spectral channels for normal-hearing listeners (dashed line with small filled symbols) or as a function of the number of electrodes used with Nucleus-22 cochlear implant listeners (filled symbols) and Clarion cochlear implant listeners (open symbols). The hatched area plots the range of performance across all 19 cochlear implant listeners. From left to right the panels present consonant recognition as a function of decreasing signal-to-noise ratio.

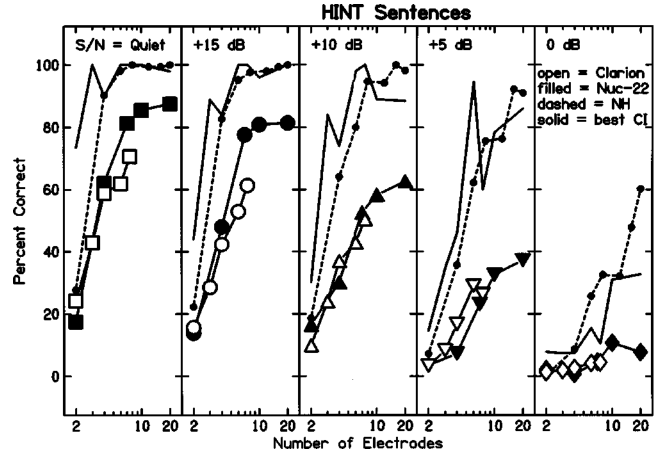

FIG. 4. Recognition of HINT sentences as a function of the number of spectral channels for normal-hearing listeners (dashed line with small filled symbols) or as a function of the number of electrodes used with Nucleus-22 cochlear implant listeners (filled symbols) and Clarion cochlear implant listeners (open symbols). The solid line plots the best performance level across all 19 cochlear implant listeners. From left to right the panels present consonant recognition as a function of decreasing signal-to-noise ratio.

the open symbols present data from subjects with the Clarion device, filled symbols present data from subjects with the Nucleus-22 device, and the dashed line with small filled symbols presents results from normal-hearing listeners with noise-band processors. Average standard deviations for the three types of listeners on the four sets of test materials are given in Table III. Note that the variability was similar for the two sets of implant listeners, while the standard deviation for normal-hearing listeners was generally about half that observed in the implant listeners. The hatched area in Figs. 1 and 2 outlines the entire range of performance across all 19 implant subjects. Tables IV and V present listener scores for all tests in quiet with their original processor.

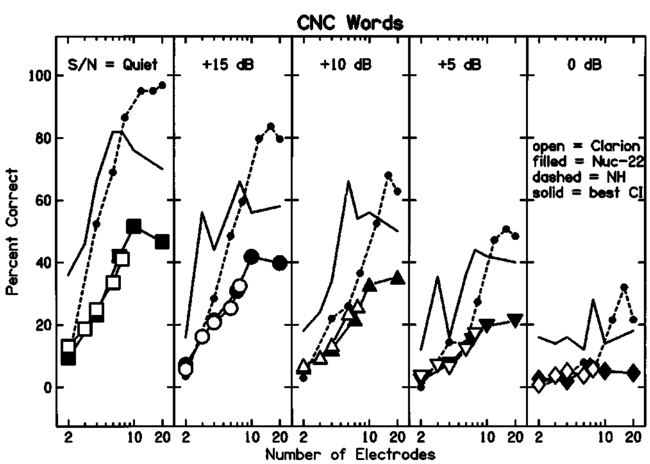

FIG. 3. Recognition of CNC words as a function of the number of spectral channels for normal-hearing listeners (dashed line with small filled symbols) or as a function of the number of electrodes used with Nucleus-22 cochlear implant listeners (filled symbols) and Clarion cochlear implant listeners (open symbols). The solid line plots the best performance level across all 19 cochlear implant listeners. From left to right the panels present consonant recognition as a function of decreasing signal-to-noise ratio.

A. Comparison of implants

For Clarion listeners, a repeated measures ANOVA revealed no difference in performance between the CIS and SAS patients (Table VI) for all numbers of electrodes and all noise levels. The two groups of Clarion listeners were then grouped together for comparison with Nucleus listeners. For this comparison the Nucleus results from 7, 10, and 20 electrodes were all grouped together and compared to Clarion results with 7 or 8 electrodes. In all other cases similar numbers of electrodes were compared across the two devices. A repeated measures ANOVA revealed no significant difference in performance for listeners with the two implants for all conditions (Table VI). Speech recognition with both the Clarion and Nucleus-22 processors improved as the number of electrodes increased (up to seven or eight) for all conditions. Even though more electrodes were available with the Nucleus-22 speech processor, performance was not significantly better than for seven or eight electrodes with the Clarion device. A repeated measures ANOVA was performed, comparing results for Nucleus CI listeners with 7, 10, and 20 electrodes (Table VII). While there was no significant difference in speech recognition for consonants or vowels, there was a marginally significant effect for words and sentences. *Post-hoc t*-tests revealed that there was a marginally significant difference between the seven- and ten-electrode results for word recognition in quiet and at +15 dB S/N ratio ($p < 0.05$), but not at other noise levels. *Post-hoc*

TABLE III. Average standard deviations (%) for each test and listener type. Standard deviations were averaged across noise conditions and number of channels.

	Normal-hearing	Clarion	Nucleus-22
Vowels	4.87	11.50	10.95
Consonants	3.58	10.81	9.67
CNC words	7.50	11.86	10.32
HINT sentences	7.24	19.74	15.21

TABLE IV. Nucleus-22 listeners' scores (%) in quiet with original processor.

Listener	Strategy	Vowels	Consonants	CNC words	HINT sentences
N3	SPEAK	53	47	46	85
N4	SPEAK	68	71	70	100
N6	SPEAK	67	60	54	92
N7	SPEAK	57	58	46	92
N9	SPEAK	72	76	76	100
N14	SPEAK	29	32	12	58
N15	SPEAK	58	57	34	89
N17	SPEAK	37	51	28	81
N18	SPEAK	75	60	44	96
N19	SPEAK	58	55	66	94
Mean score		57	57	48	89

tests revealed no significant differences between seven and ten electrodes for sentences at any noise level.

B. Comparison of acoustic and electric hearing

The performance of NH listeners was significantly better than CI listeners for all noise conditions (Table VI). In addition, scores of the NH listeners continued to increase up to 20 channels with similar signal processing conditions. For consonant recognition (Fig. 1), NH listeners scored consistently higher than the best CI listeners for all numbers of electrodes/bands, particularly at high signal-to-noise levels. For vowel recognition (Fig. 2), performance by the best CI listeners (top edge of the hatched area) was similar to NH listeners, but only up to eight electrodes. As the number of electrodes was increased beyond eight, performance for the CI listeners remained relatively constant, while for NH listeners, performance continued to increase with the number of bands. For CNC word recognition (Fig. 3) and HINT sentence recognition (Fig. 4) only the best performance by CI patients is presented because the lowest performance level was near zero for all conditions. As with the phoneme results, the best performance level with cochlear implants was similar to that of NH listeners with the same processing, up to seven to eight channels/electrodes. As the number of channels/electrodes was increased above seven to eight, word and sentence recognition continued to increase in normal-hearing listeners. The line representing the best implant score is somewhat erratic because it represents a single score. With a small number of channels/electrodes and at high noise levels the best implant listeners appear to have higher scores than NH listeners for the same conditions. This might be due to the greater experience of implant listeners in such difficult listening conditions. In their everyday listening experience implant listeners must constantly reconstruct linguistic information from partially received phonemic fragments, whereas NH listeners only face such a challenging reconstruction problem infrequently.

TABLE V. Clarion listeners' scores (%) in quiet with original processor.

Listener	Strategy	Vowels	Consonants	CNC words	Hint sentences
C1	CIS	46	52	20	87
C3	CIS	57	60	64	89
C4	CIS	28	25	28	40
C5	CIS	78	75	82	100
C9	CIS	26	33	22	60
C2	SAS	58	65	48	91
C6	SAS	77	55	48	96
C7	SAS	45	36	24	40
C8	SAS	43	42	34	34
Mean score		51	49	41	75

TABLE VI. Repeated measures ANOVA F-values between subjects (df $=1$ for all F values).

Test	CIS-SAS	Nucleus-Clarion	NH-CI	Better-poorer listeners
Vowels	0.025	0.812	13.981[a]	71.579[a]
Consonants	0.003	1.748	32.840[a]	25.210[a]
Words	0.215	0.065	16.776[a]	28.252[a]
Sentences	0.419	0.165	25.598[a]	58.821[a]

[a] $p<0.05$.

C. Comparison of good and poor implant scores

One interesting feature of the results can be observed by comparing the top and bottom borders of the hatched area in Figs. 1 and 2. The hatched area represents the overall range of scores obtained from all 19 implant listeners. The top and bottom edges of the hatched area do not necessarily represent the scores from single implant listeners, although they are representative of the performance curves from the better and poorer implant listeners. The better-performing implanted listeners improved as the number of electrodes was increased up to seven, while the poorer-performing listeners showed little increase in performance as the number of electrodes was increased above four. This was particularly true in tests that rely more heavily on spectral cues (such as vowel recognition) and at low S/N ratios. A repeated measures ANOVA was performed between the two groups of performers, revealing a statistically significant difference for all tests (Table VI). This result suggests that implant listeners with better speech recognition are able to utilize more channels of spectral information than those with poor speech recognition. To test this hypothesis, consonant confusion matrices were analyzed for the two implant listeners with the best scores and the two implant listeners with the poorest scores.

Consonant recognition was analyzed into the traditional production-based categories of voicing, manner of articulation, and place of articulation (Miller and Nicely, 1955). The percent correct for each of these features is presented in Fig. 5 for the two best users and the two poorest users of each device and for NH listeners with similar processing. The results show that the better implant users were able to utilize all three categories of cues better than the poorer implant users. A repeated measures ANOVA was performed on the feature scores, revealing a statistically significant difference between the better and poorer hearing listeners (voicing: $F=33.728$, $p<0.001$, manner: $F=11.595$, $p<0.05$, and place: $F=14.279$, $p<0.01$). At high signal-to-noise ratios the number of electrodes did not affect the reception of voicing. However, at poor signal-to-noise ratios the percent correct on voicing increased with the number of electrodes up to four. This may be due to the noise interfering with the per-

TABLE VII. Repeated measures ANOVA F-values within subjects for Clarion, Nucleus-22 (all electrode combinations), Nucleus-22 (7-, 10-, and 20-electrode maps only), all cochlear implant listeners (Nucleus-22 and Clarion combined), and normal-hearing listeners.

Electrode factor	Clarion	Nucleus-22	Nucleus-22 (7-, 10-, 20-electrode maps only)	CI listeners	Normal-hearing listeners
Vowels	20.648[b]	41.633[b]	3.202	53.544[b]	159.768[b]
Consonants	15.543[b]	27.603[b]	1.147	59.438[b]	222.124[b]
Words	13.383[b]	38.270[b]	7.461[a]	45.854[b]	56.918[b]
Sentences	11.912[b]	59.982[b]	5.288[a]	56.632[b]	247.631[b]
Noise factor					
Vowels	21.966[b]	34.161[b]	40.066[b]	36.400[b]	4.448[b]
Consonants	62.346[b]	81.310[b]	60.389[b]	146.019[b]	4.184[b]
Words	17.487[b]	52.858[b]	54.645[b]	40.335[b]	13.531[b]
Sentences	36.937[b]	132.997[b]	141.768[b]	118.740[b]	6.825[b]
Electrode–noise interaction					
Vowels	2.355[a]	6.311[b]	1.965	9.428[b]	120.085[b]
Consonants	2.212[a]	0.791[a]	0.639	2.594[a]	673.979[b]
Words	4.924[b]	6.830[b]	1.681	9.637[b]	633.295[b]
Sentences	4.521[b]	12.693[b]	0.438	27.952[b]	364.443[b]

[a] $p<0.05$.
[b] $p<0.001$.

ception of the temporal cues for voicing, which should not require much spectral information. When temporal cues are masked by noise, voicing can be conveyed by spectral cues (e.g., the ratio of energy above and below 1500 Hz), but multiple electrodes are required to provide this spectral information. A similar pattern of performance was observed for the reception of manner cues. Percent correct on the place of articulation increased as the number of electrodes increased

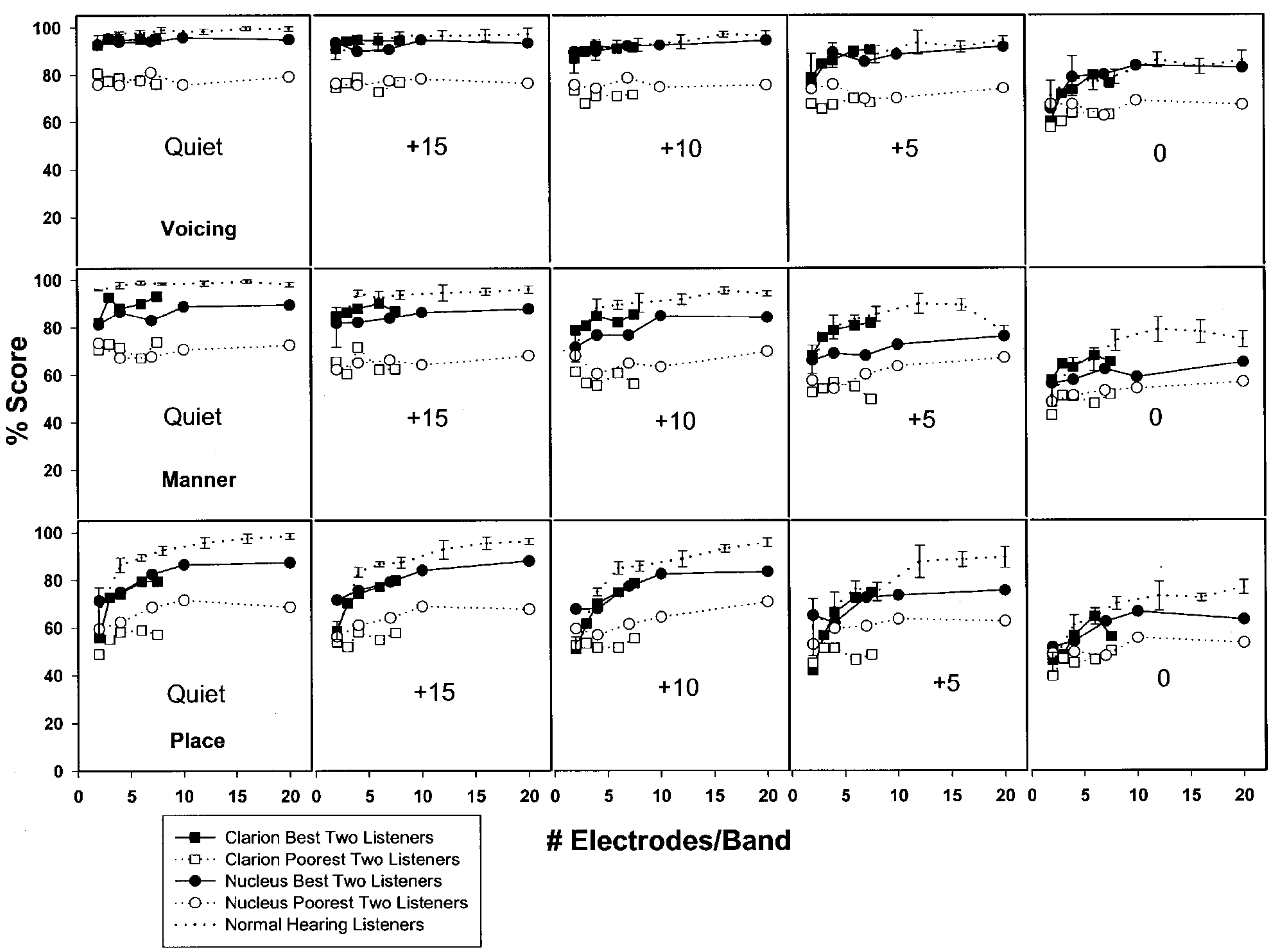

FIG. 5. Percent correct on the consonant features of voicing (top row), manner (middle row), and place of articulation (lower row) as a function of the number of electrodes or number of spectral bands. Dotted lines with error bars represent the range of results from normal-hearing listeners. Filled symbols present the average of the two best scores from CI listeners with each device, while open symbols present the average of the two poorest scores from CI listeners with each device.

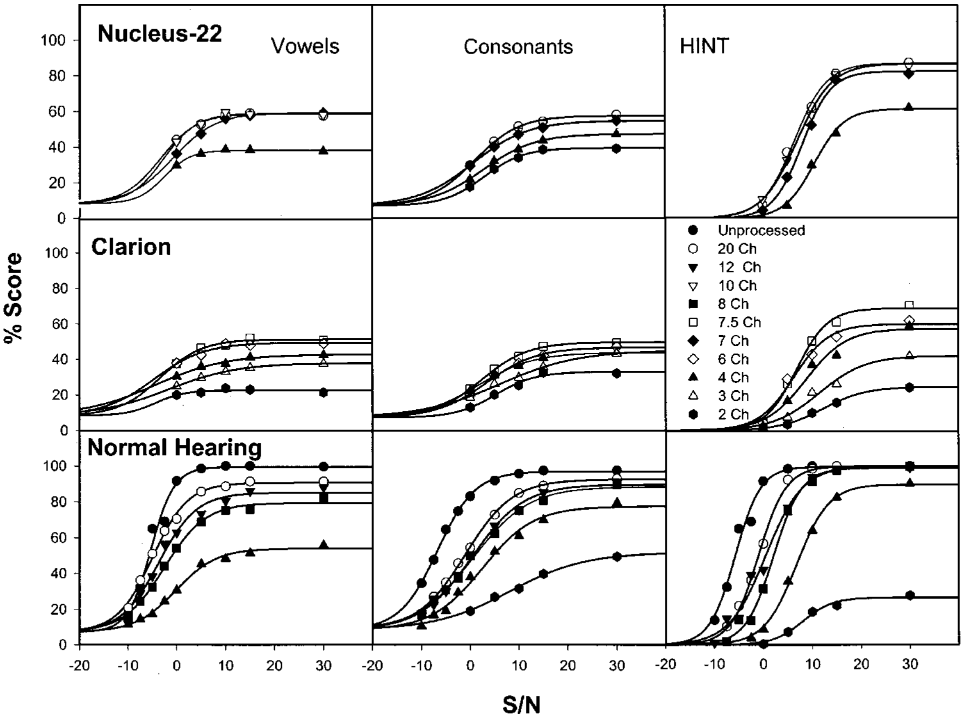

FIG. 6. Recognition functions as a function of signal-to-noise ratio for NH and CI listeners for consonant, vowel, and sentence recognition. Lines plot the fit of a sigmoidal model [Eq. (1)] to the data. Parameters of the model from these fits, as well as the rms error of each fit, are presented in Tables VIII–X.

at all noise levels (but only increased up to seven to ten electrodes for implant listeners).

In this study, the better listeners with the Nucleus-22 device were receiving similar amounts of information on voicing, manner, and place as the better listeners with the Clarion device (both CIS and SAS). In addition, scores for the best implant listeners were similar to that of the normal-hearing listeners with the same number of processing channels. Scores from poor-performing listeners with both implant devices were also similar to each other, indicating that the electrode design and speech processor strategy were not the primary factors affecting speech recognition or the pattern of performance across implant listeners. Note that the poorer-performing implant listeners were only receiving 80% correct on voicing and 70% correct on manner cues, regardless of the number of electrodes. This is an unusual result because voicing and manner cues are thought to be conveyed primarily by temporal information and should be easily available to implant listeners. The better-performing implant listeners and NH listeners received essentially 100% of the voicing and manner information with only two channels of spectral information. This is consistent with a previous result by Shannon *et al.* (1995). Yet the poorer-performing implant listeners never achieved this level of performance even with the maximum number of electrodes.

D. Performance-intensity functions

The vowel, consonant, and sentence recognition data from Figs. 1–4 are replotted in Fig. 6 as a function of SN ratio. Data from each condition was fit with a simple three-parameter sigmoidal model (Boothroyd *et al.*, 1996):

$$\%C = P_0 + (Q - P_0)/(1 + e^{-\beta(x-\mathrm{PRT})}), \tag{1}$$

where P_0 is the chance performance level, Q is the percent correct in quiet, β is related to the slope of the function, x is the level of the noise in dB, and PRT is the phoneme recognition threshold in dB, which is the S/N ratio at which the performance falls to 50% of the level in quiet. The values of Q, β, and PRT and the standard error of the fit are presented in Tables VIII–X for consonant, vowel, and sentence recognition, respectively. As can be seen in Fig. 6 and Tables VIII–X, this function provided excellent fits to all curves of performance as a function of signal-to-noise ratio. In general, as the number of spectral channels/electrodes decreased, the level of performance in quiet decreased, and the PRT increased. The slopes of the functions were relatively constant across conditions, decreasing slightly for the two-channel condition. The parameter values obtained for the NH listeners were similar to those obtained under similar conditions, but with different speech materials, by Fu *et al.* (1998).

Figure 7 plots the PRT as a function of the number of spectral channels (or number of electrodes) for consonant, vowel, and sentence recognition. The PRTs for NH listeners from the present study (filled squares) are comparable to PRTs from NH listeners in a previous study (Fu *et al.*, 1998). In both studies, the PRT decreased linearly as a function of the logarithm of the number of spectral channels. Regression slopes are presented in Table XI. Slopes of linear regression fits to the NH data are −2.57, −2.44, and −2.87 dB/doubling of the number of channels for consonants, vowels, and sentences, respectively (Table XI). In contrast, the PRT for CI listeners, except for the Clarion listeners with HINT sentences, changed little as the number of electrodes was increased.

IV. DISCUSSION

A. Comparison of implant devices and processing strategies

A key result in the present study is that there was no significant difference in speech recognition performance be-

TABLE VIII. Sigmoidal model parameters for consonant recognition. (PRT=phoneme recognition threshold, β=related to the slope of the function, and Q=percent correct in quiet.)

Listeners	Channels/Electrodes	PRT (dB)	Q (%)	β	Standard error of fit (%)
Nucleus-22	2	2.90±0.31	39.6±0.6	0.21±0.38	0.68
	4	2.70±0.31	47.5±0.7	0.20±0.36	0.64
	7	0.59±0.20	54.9±0.5	0.17±0.28	0.45
	10	0.90±0.09	56.0±0.2	0.20±0.16	0.24
	20	0.90±0.29	57.6±0.8	0.21±0.38	0.85
Clarion	2	5.23±1.04	33.0±1.7	0.24±0.95	1.75
	3	6.16±1.14	45.1±2.0	0.12±1.26	1.22
	4	1.75±0.34	43.9±0.6	0.17±0.44	0.61
	6	3.42±0.58	46.7±1.3	0.19±0.64	1.23
	7/8	2.33±0.29	49.7±0.7	0.42±0.35	0.71
NH	2	7.81±1.06	52.0±2.2	0.13±1.05	1.29
	4	2.30±0.73	77.7±2.6	0.18±0.53	2.78
	6	1.16±0.34	84.7±1.3	0.17±0.26	1.37
	8	0.30±0.45	88.6±1.8	0.17±0.37	1.92
	12	0.13±0.31	89.9±1.3	0.18±0.25	1.48
	16	−0.89±0.44	93.0±1.9	0.18±0.38	2.21
	20	−1.21±0.22	92.7±0.6	0.20±0.20	1.25
	Unprocessed	−6.86±0.12	96.9±0.6	0.26±0.14	1.06

tween Nucleus and Clarion implant systems either as a function of the signal-to-noise ratio or as the number of electrodes was varied. For each of these conditions the range of speech recognition performance was wide, but the range of performance and average scores were not different between Nucleus and Clarion devices, or between the SAS and CIS strategies in Clarion patients. Indeed, in our limited sample even the average scores of the two best implant listeners and the two poorest implant listeners were similar between the Nucleus and Clarion systems. In this sample of ten Nucleus and nine Clarion implant listeners there were no differences in performance in spite of the large differences between the systems tested (i.e., electrode design and placement, analog versus pulsatile stimulation, SPEAK versus SAS versus CIS, fast versus slow stimulation rate). In contrast, the number of electrodes had a large and significant effect on all devices and speech processing strategies.

B. Effect of S/N ratio

Figure 6 plots the speech recognition scores as a function of the S/N ratio for consonants, vowels, and sentences. Note that all curves are well fit by the three-parameter sigmoidal function of Eq. (1). The standard error of the fits is less than 3% for most curves (Tables VIII–X). The slopes of the normalized functions are similar for high numbers of channels, but are shallower when the number of channels is small, as has been documented previously (Boothroyd *et al.*, 1996; Fu *et al.*, 1998).

TABLE IX. Sigmoidal model parameters for vowel recognition. (PRT=phoneme recognition threshold, β =related to the slope of the function, and Q=percent correct in quiet.)

Listeners	Channels/Electrodes	PRT (dB)	Q (%)	β	Standard error of fit (%)
Nucleus-22	2	⋯	⋯	⋯	
	4	−2.50±0.70	38.4±0.5	0.36±0.69	0.67
	7	−2.06±0.39	59.6±0.8	0.21±0.50	0.88
	10	−2.60±1.03	58.9±1.3	0.29±1.09	1.45
	20	−3.30±1.11	58.7±1.1	0.26±1.08	1.75
Clarion	2	−4.33±7.17	22.5±0.9	0.33±4.92	1.32
	3	−1.63±0.04	37.9±0.0	0.14±0.06	0.04
	4	−4.39±1.24	42.8±1.0	0.14±1.54	0.87
	6	−4.64±2.01	49.2±1.4	0.19±1.95	1.56
	7/8	−3.30±1.25	51.3±1.2	0.23±1.30	1.45
NH	2	⋯	⋯	⋯	⋯
	4	0.02±0.43	53.7±1.1	0.25±0.35	1.57
	6	−1.95±0.29	68.8±0.9	0.22±0.27	1.33
	8	−2.39±0.29	79.8±1.2	0.24±0.27	1.76
	12	−3.56±0.47	84.9±2.2	0.26±0.46	2.35
	16	−4.81±0.55	88.6±2.6	0.26±0.58	4.38
	20	−4.86±0.27	90.6±1.4	0.28±0.28	2.35
	Unprocessed	−6.75±0.13	96.9±0.6	0.26±0.15	1.10

TABLE X. Sigmoidal model parameters for sentence recognition. (PRT=phoneme recognition threshold, β =related to the slope of the function, and Q=percent correct in quiet.)

Listeners	Channels/Electrodes	PRT (dB)	Q (%)	β	Standard error of fit (%)
Nucleus-22	2	⋯	⋯	⋯	⋯
	4	10.05 ± 0.50	61.5 ± 2.4	0.33 ± 0.40	2.42
	7	8.05 ± 0.37	82.6 ± 2.3	0.33 ± 0.30	2.51
	10	7.02 ± 0.43	86.8 ± 2.5	0.28 ± 0.35	2.64
	20	6.47 ± 0.53	87.1 ± 3.2	0.30 ± 0.44	3.55
Clarion	2	12.22 ± 0.51	24.5 ± 0.8	0.23 ± 0.43	0.70
	3	11.42 ± 1.55	42.0 ± 4.0	0.21 ± 1.30	3.39
	4	8.73 ± 1.44	57.3 ± 5.3	0.24 ± 1.18	5.17
	6	6.27 ± 1.17	59.9 ± 4.6	0.28 ± 1.00	5.04
	7/8	6.86 ± 0.59	68.8 ± 2.9	0.31 ± 0.49	3.19
NH	2	8.01 ± 0.95	26.6 ± 1.8	0.11 ± 0.76	2.05
	4	6.73 ± 0.29	89.8 ± 1.8	0.11 ± 0.76	2.05
	6	3.52 ± 0.46	96.6 ± 2.82	0.31 ± 0.33	3.65
	8	2.03 ± 0.38	99.1 ± 2.3	0.35 ± 0.27	3.45
	12	0.46 ± 0.61	100.0 ± 3.5	0.26 ± 0.48	4.87
	16	-0.62 ± 0.67	100.0 ± 4.3	0.31 ± 0.57	6.51
	20	-1.01 ± 0.19	100.0 ± 1.2	0.35 ± 0.17	2.06
	Unprocessed	-5.75 ± 0.36	99.7 ± 2.5	0.38 ± 0.35	4.92

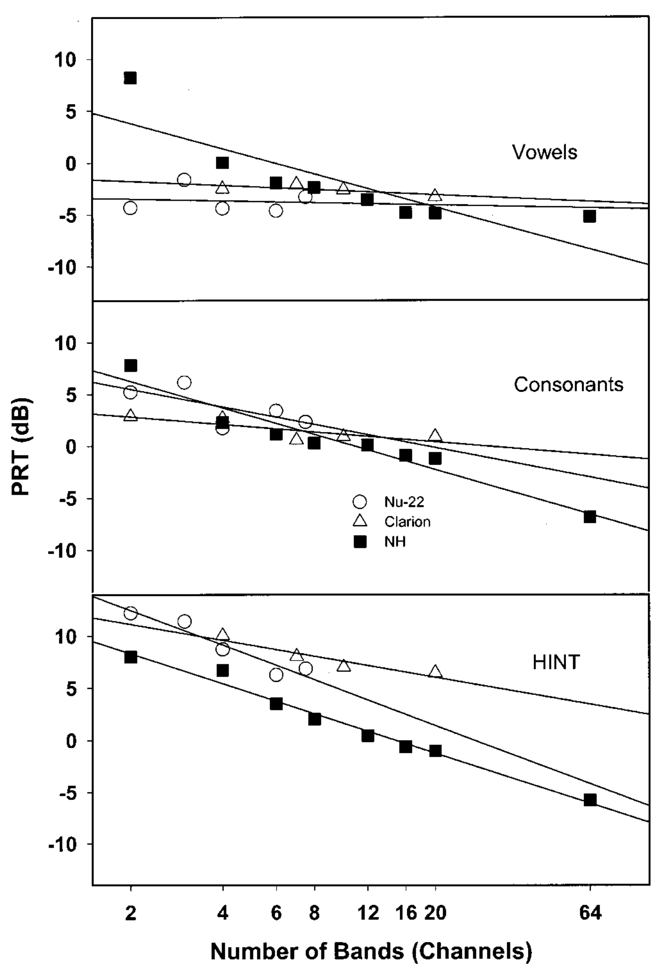

FIG. 7. Phoneme recognition threshold (PRT) as a function of the number of electrodes or bands for consonant (top), vowel (middle), or sentence (bottom) recognition. Filled symbols present data from NH listeners, while open symbols present data from CI listeners. Regression lines are fit to each set of data, with the parameters and fits listed in Table XI.

At high S/N ratios, sentence recognition is high for all number of channels greater than three. At low S/N ratios, however, a reduction in the number of channels is equivalent to reducing the S/N ratio. The best performing CI listeners need a 5–10 dB better S/N ratio to obtain performance equivalent to normal-hearing listeners using 20 channels, indicating that even the best CI listeners are using the equivalent of about 8 channels. One implication of this result is that implant listeners would be able to understand speech much better in noise if we could increase the number of spectral channels effectively used. Because the sentence recognition function has a slope of 6%–10%/dB, a 5-dB difference in S/N ratio could potentially produce a 30%–50% improvement in sentence recognition.

C. Effect of the number of electrodes

The results in quiet basically replicate the overall pattern of results of Fishman *et al.* (1997), who also found that speech recognition for Nucleus-22 listeners improved with the number of electrodes. However, Fishman *et al.* found no further improvement as the number of electrodes increased beyond four. The present results show a similar pattern of results, although speech recognition continued to improve up to seven electrodes for vowel and consonant recognition and up to ten electrodes in Nucleus listeners for word and sen-

TABLE XI. Slopes (dB/doubling of functions from Fig. 7) relating PRT and the log number of channels from the present study and Fu *et al.* (1998).

Listeners	Consonants Slope	Vowels Slope	Sentences Slope
Nucleus-22	-0.73	-0.40	-1.55
Clarion	-1.69	-0.17	-3.33
Normal-hearing	-2.57	-2.44	-2.87
Fu *et al.* (1988) (NH)	-2.90	-2.35	⋯

tence recognition. This improvement up to ten electrodes appeared to be more evident at medium noise levels (see Figs. 3 and 4).

In this study there did appear to be a slight difference in the pattern of performance between the better implant users and the poorer users. The poorer users did not improve in performance as the number of electrodes was increased beyond three or four. This result is different from the result observed by Fishman *et al.* (1997) where both good and poor implant users improved only up to four electrodes. A possible reason for this difference is that the poorer performers in this study had lower scores than those in the Fishman *et al.* study. The larger range of performance in the present study may have accentuated a real difference between good and poor implant users, with this study showing that poor performance was limited by the ability to use information from more than four electrodes. It is possible that the use of multiple electrodes was limited by electrode interactions.

Speech recognition with the Clarion device increased as the number of electrodes increased up to the maximum number available (eight for CIS, seven for SAS). It would be interesting to see if adding more channels to the Clarion strategies resulted in an increase in performance, or whether the performance would reach a plateau at seven or eight electrodes similar to the SPEAK results. The newest version of the Clarion implant (the C-II) will allow 16 independent channels of stimulation at high stimulation rates. The present experiment should be repeated with that new device to see if the performance level asymptotes at eight electrodes or not.

In listeners with the Nucleus-22 device utilizing the SPEAK processing strategy, speech recognition performance increased as the number of electrodes increased, but only up to seven to ten electrodes. There were no significant differences in performance for 7, 10, or 20 electrodes for vowel and consonant recognition, although there was a significant improvement from 7 to 10 electrodes for word and sentence recognition. Normal-hearing listeners, in contrast, continued to improve in speech recognition as the number of spectral bands was increased, at all noise levels.

One of the most puzzling aspects of the present data is that even the best CI performance appears to be limited to the equivalent of seven to ten spectral channels. In the Fishman *et al.* (1997) study, an asymptote in performance was observed with four electrodes for consonants and CUNY sentences and seven electrodes for the more difficult tests of vowels and NU6 words (although the increase in performance from four to seven electrodes did not achieve statistical significance). Our results showed an asymptote of seven electrodes for consonants and vowels and ten electrodes for words and sentences. One explanation for the difference may be due to the more difficult test materials used in this study for vowels and consonants (male and female multitalkers), compared to the Fishman *et al.* study (one male talker). However, this difference in test difficulty does not apply to words and sentences. Another possible reason for the difference is that in the present experiment the number of subjects and distribution of scores may have provided sufficient statistical power to show a significant improvement from four to seven electrodes for vowels and consonants and from seven to ten electrodes for words and sentences.

Let us consider three hypotheses to explain the differences between CI and NH performance, in particular what factors might limit performance in CI listeners to seven to ten channels.

D. Hypothesis 1: Stimulation rate is the primary factor limiting performance

Consider the hypothesis that the limitation of implant listeners to seven to ten channels of spectral information is due to the relatively low pulse rate/electrode in the SPEAK processing strategy. If this hypothesis is correct, then a SPEAK or CIS system with a faster pulse rate per electrode might show an improvement in performance as the number of electrodes is increased above seven. The CIS strategy implemented in the Nucleus-24 device allows stimulation of up to 12 electrodes at stimulation rates of up to 1200 pps/electrode. The ACE strategy, which is a hybrid of SPEAK and CIS strategies allows stimulation of up to 20 electrodes out of 22 at rates of up to 720 pps/electrode. These stimulation rates are considerably higher than rates allowed by the Nucleus-22 SPEAK processor. However, preliminary data (Arndt *et al.*, 1999) indicate no difference between the 22 and 24 systems in the average level of performance, even with the full number of electrodes and higher stimulation rates, so it is unlikely that patients with the 24 systems are improving in performance beyond seven electrodes. More data is needed to confirm this observation. If this preliminary result is confirmed, it would suggest that stimulation rate is not the primary factor limiting the number of usable channels of spectral information. Because of the high variability across CI listeners, this comparison should be done within subjects.

Several groups (Wilson *et al.*, 1988; Rubinstein *et al.*, 1999; Chatterjee and Robert, 2001) have recently suggested that very high stimulation rates (>4 kHz/electrode) might aid channel independence by producing stochastic neural firing on each electrode. The hypothesis is that synchronous firing across electrodes, as probably occurs with low-rate stimulation, can cause the entire pattern of firing to group together as a single ‘‘auditory object.’’ This grouping may not allow the information from each electrode to be usable independently. Stimulation rates high enough to introduce stochastic firing near each electrode may overcome this ‘‘forced’’ grouping and allow the information on each electrode to contribute independently. Special interfaces and implant devices are necessary to test this hypothesis.

E. Hypothesis 2: Electrode interaction is the primary factor limiting performance

It seems reasonable to assume that interaction between electrodes would reduce the effective tonotopic selectivity of a multichannel implant and thus could limit the listeners ability to understand speech. Cochlear current spread, producing interaction between electrodes, may limit spectral resolution in cochlear implants. Physically, electrical current spread should be greater at higher stimulation levels and for monopolar stimulation modes. However, some researchers have noted better electrode discrimination at higher overall levels, indicating that increased current spread in the cochlea does not necessarily lead to poorer electrode discrimination (McKay *et al.*, 1999; Pfingst *et al.*, 1999). In addition, speech recognition performance may even improve with stimulation level (Skinner *et al.*, 1997). Several studies have shown that monopolar stimulation mode, which should produce a broad current distribution, can provide the same or better speech perception ability than bipolar stimulation modes (Kileny *et al.*, 1998; Zwolan *et al.*, 1996).

Three of the implant listeners in the present study (N3, N4, and N7) were also subjects in two previous psychophysical studies of electrode interaction (Chatterjee and Shannon, 1998; Hanekom and Shannon, 1998). Chatterjee and Shannon (1998) measured forward masking patterns across the electrode array as a measure of the spread of excitation in a cochlear implant. They observed that excitation patterns measured in cochlear implants were broader than similar measures in acoustic hearing, but saw no widening of the patterns with level. Hanekom and Shannon (1998) measured electrode interaction using gap detection, also with listeners N3, N4, and N7. In both studies, listener N3 had the most electrode interaction and poorest speech recognition and N4 had the least interaction and the highest speech recognition. Listener N3 also showed more changes in the pattern of electrode interaction as stimulation parameters were changed than N4 and N7. While there seems to be a rough association between electrode interaction and speech recognition, the relation is does not appear to be strong one. And it does not appear that electrode interaction is the limiting factor for listeners like N4, who show little electrode interaction yet show the same asymptote in speech recognition with seven electrodes as other listeners.

Observe the difference between the upper and lower edge of the hatched area in Figs. 1 and 2. The upper edge, which represents the best implant scores across all 19 listeners increases in performance up to seven or eight channels. The lower edge, representing the poorest scores across all implant listeners, does not increase substantially from 4 to 20 electrodes. Thus, consistent with the electrode interaction hypothesis, poor CI speech recognition is limited to three or four effective spectral channels, while good implant speech recognition improves with the number of electrodes, up to seven or eight.

Several other studies have shown a modest correlation between electrode interaction and speech recognition performance (Nelson *et al.*, 1995; Hanekom and Shannon, 1996; Throckmorton and Collins, 1999; Zwollan *et al.*, 1997; Collins *et al.*, 1997; Donaldson and Nelson, 2000; Henry *et al.*, 2000). In general, cochlear implant listeners with more electrode interaction contain lower speech recognition scores, but only a small portion of the variance is accounted for by the electrode interaction.

If electrode interactions are limiting performance on the top end of implant performance with seven to ten channels, then the poorer users may have an increased amount of electrode interaction that limits their performance even further—not allowing performance to improve beyond the three- to four-channel level, no matter how many electrodes are used. Inspection of the results in Figs. 1 and 2 shows that the implant listeners who are poor at speech recognition did not improve as the number of electrodes was increased above three or four. If this is the case, then it is of utmost importance to discover the cause of the electrode interactions and either correct this problem in the signal processing or with new electrode designs.

F. Hypothesis 3: Warping in the spectral-tonotopic mapping is the primary factor limiting performance

Another possible cause of the limitation in the use of all channels is the presence of distortion in the representation of the spectral information. Fu and Shannon (1999a) found that speech recognition was reduced when the spectral information was represented at cochlear locations that were shifted either apically or basally from their normal location. They found the same pattern of results for 4, 8, and 16 bands of spectral resolution, indicating that higher levels of spectral resolution did not mitigate the negative effects of a frequency-place shift. Shannon *et al.* (1998) found that speech recognition was reduced when the tonotopic distribution of spectral information was warped nonlinearly from its normal acoustic mapping. Both studies observed that a warping in the tonotopic distribution could not only result in a reduced number of effective channels of spectral information, but could also reduce the reception of what are thought to be primarily temporal cues in speech, like voicing and manner. Fu (1997) and Shannon *et al.* (1998) saw significant reductions in voicing, manner, and place information received on consonants when the tonotopic mapping was distorted. A similar pattern was observed in the poorer implant listeners in this study, i.e., their reception of voicing and manner were significantly poorer than in the implant users with better performance. In normal-hearing listeners nearly 100% of the voicing and manner information in consonants is received for all processors with two or more channels [Shannon *et al.* (1995) and Fig. 5]. Compare the reception of voicing and manner cues in Fig. 5 between the normal-hearing listeners and the two groups of implant listeners. The implant listeners with poor performance received less voicing and manner information with 20 electrodes than the NH and better-performing implant listeners received with only 2 electrodes. Thus, a reduction in the reception of voicing and manner cues indicates more than simply a loss of the number of effective channels of spectral information. The reduction in speech recognition due to frequency-place distortions appears to be independent of spectral resolution, and so would exacerbate the reduction due to a reduced number of channels of spectral information. Based on the Fu and Shannon (1999a) results on tonotopic shifting and the Fu (1997) and Shannon *et al.* (1998) results on frequency-place warping, we suggest that overall poor reception of voicing and manner cues could indicate the presence of a shift or warping in the frequency-place mapping in those patients.

V. CONCLUSIONS

Speech recognition was similar for the Clarion and Nucleus-22 cochlear implant listeners in this study both as a function of noise level and as function of number of electrodes. Speech recognition improved significantly as a function of the number of electrodes up to seven to ten electrodes. No improvement was observed in speech recognition as the number of electrodes was increased from 7 to 20 for vowels and consonant recognition and no improvement was observed as the number of electrodes was increased from 10 to 20 for word and sentence recognition. For the limited sample size of the present study, this pattern of results suggests that the number of electrodes is a more important factor in implant performance than the differences in design between the two implants. Comparison to normal-hearing listeners with similar processing suggests that some cochlear implant listeners can fully utilize the spectral information provided (up to eight electrodes), but others do not. The relatively small improvement for the poorer implant performers as the number of spectral channels was increased suggests that these individuals are not able to utilize the spectral information provided. Even the implant listeners with the best speech recognition appeared to be unable to utilize more than seven to ten channels of spectral information, no matter how many channels of information are presented. The reason for this limitation is not clear. We speculate that this limitation may be due to electrode interactions, and to possible tonotopic shifts and warping in the frequency-to-place mapping of spectral information.

ACKNOWLEDGMENTS

We thank the cochlear implant and normal-hearing listeners for their patience and persistence in these experiments. A special thanks goes to Dr. Qian-Jie Fu for his insight and to Mark Robert for his technical help. We also appreciate the helpful comments of two anonymous reviewers. Funding for this research was provided by NIH Grant No. R01-DC-01526 and NIH Contract No. N01-DC-92100.

Arndt, P., Staller, S., Arcaroli, J., Hines, A., and Ebinger, K. (**1999**). *Within subject comparison of advanced coding strategies in the Nucleus 24 cochlear implant* (Cochlear Corporation, Englewood, CO).

Baer, T., and Moore, B. C. J. (**1993**). ''Effects of spectral smearing on the intelligibility of sentences in noise,'' J. Acoust. Soc. Am. **94**, 1229–1241.

Baer, T., and Moore, B. C. J. (**1994**). ''Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech,'' J. Acoust. Soc. Am. **95**, 2277–2280.

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

Breeuwer, M., and Plomp, R. (**1984**). ''Speechreading supplemented with frequency-selective sound-pressure information,'' J. Acoust. Soc. Am. **76**, 686–691.

Chatterjee, M., and Robert, M. (**2001**). ''Noise enhances modulation sensitivity in cochlear implant listeners: Stochastic resonance in a prosthetic sensory system?,'' J. Assoc. Res. Otolaryngol. **2**(2) (in press).

Chatterjee, M., and Shannon, R. V. (**1998**). ''Forward masked excitation patterns in multielectrode cochlear implants,'' J. Acoust. Soc. Am. **103**, 2565–2572.

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

Cochlear Corporation (**1995**). *Audiologist Handbook* (Englewood, CO).

Collins, L. M., Zwolan, T. A., and Wakefield, G. H. (**1997**). ''Comparison of electrode discrimination pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects,'' J. Acoust. Soc. Am. **101**, 440–455.

Collins, L., Zwolan, T., O'Neill, J., and Wakefield, G. (**1994**). ''Analysis of electrode pair confusions and implications for speech recognition in cochlear implant listeners,'' in *Abstracts of the Seventeenth Midwinter Meeting, Association for Research in Otolaryngology*, edited by D. Lim, p. 161.

Donaldson, G. S., and Nelson, D. A. (**2000**). ''Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies,'' J. Acoust. Soc. Am. **107**, 1645–1658.

Dorman, M. F., and Loizou, P. C. (**1997**). ''Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants,'' Am. J. Otol. Suppl. **18**, S113–S114.

Dorman, M. F., and Loizou, P. C. (**1998**). ''Identification of consonants and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normal-hearing subjects using simulations processors with two to nine channels,'' Ear Hear. **19**, 162–166.

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

Dorman, M. F., Dankowski, K., McCandless, G., and Smith, L. M. (**1989**). ''Consonant recognition as a function of the number of channels of stimulation by patients who use the Symbion cochlear implant,'' Ear Hear. **10**, 288–291.

Dubno, J. R., and Dorman, M. F. (**1987**). ''Effects of spectral flattening on vowel identification,'' J. Acoust. Soc. Am. **82**, 1503–1511.

Eddington, D. K., Rabinowitz, W. R., Tierney, J., Noel, V., and Whearty, M. (**1997**). ''Speech Processors for Auditory Prostheses,'' 8th Quarterly Progress Report, NIH Contract N01-DC-6-2100.

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

Fu, Q.-J. (**1997**). ''Speech Perception in Acoustic and Electric Hearing,'' doctoral dissertation, University of Southern California.

Fu, Q.-J., and Shannon, R. V. (**1999a**). ''Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing,'' J. Acoust. Soc. Am. **105**, 1889–1900.

Fu, Q.-J., and Shannon, R. V. (**1999b**). ''Recognition of spectrally degraded speech in noise with nonlinear amplitude mapping,'' in *Proceedings of the 1999 IEEE Conference on Acoustics, Speech and Signal Processing*, Vol. 1, pp. 369–372.

Fu, Q. J., Shannon, R. V., and Wang, X. (**1998**). ''Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing,'' J. Acoust. Soc. Am. **104**, 3586–3596.

Geier, L. V., and Norton, S. (**1992**). ''The effect of limiting the number of Nucleus 22 cochlear implant electrodes programmed on speech perception,'' Ear Hear. **13**, 340–348.

Greenwood, D. D. (**1990**). ''A cochlear frequency-position function for several species—29 years later,'' J. Acoust. Soc. Am. **87**, 2592–2605.

Hanekom, J. J., and Shannon, R. V. (**1996**). ''Place pitch discrimination and speech recognition in cochlear implant users,'' S. African J. Com. Dis. **43**, 27–40.

Hanekom, J. J., and Shannon, R. V. (**1998**). ''Gap detection as a measure of electrode interaction in cochlear implants,'' J. Acoust. Soc. Am. **104**, 2372–2384.

Henry, B. A., McKay, C. M., McDermott, H. J., and Clark, G. M. (**2000**). ''The relationship between speech perception and electrode discrimination in cochlear implantees,'' J. Acoust. Soc. Am. **108**, 1269–1280.

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

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

Holmes, A., Kemker, F. J., and Merwin, G. (**1987**). ''The effects of varying the number of cochlear implant electrodes on speech perception,'' Am. J. Otol. **8**, 240–246.

Horst, J. W. (**1987**). ''Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects,'' J. Acoust. Soc. Am. **82**, 874–885.

House Ear Institute and Cochlear Corporation (**1996**). Minimum Speech Test Battery for Adult Cochlear Implant Users CD.

Kileny, P., Zimmerman-Phillips, S., Zwolan, T., and Kemink, J. (**1992**). ''Effects of channel number and place of stimulation on performance with the cochlear corporation multichannel implant,'' Am. J. Otol. **13**, 117–123.

Kileny, P. R., Zwolan, T. A., Telian, S. A., and Boerst, A. (**1998**). ''Performance with the 20+2L lateral wall cochlear implant,'' Am. J. Otol. **19**, 313–319.

Lawson, D., Wilson, B., and Finley, C. (**1993**). ''New processing strategies for multichannel cochlear prosthesis,'' Prog. Brain Res. **97**, 313–321.

Lawson, D., Wilson, B. S., Zerbi, M., and Finley, C. C. (**1996**). ''Speech Processors of auditory prostheses,'' Third Quarterly Progress Report, NIH Contract NO1-DC-5-2103.

McDermott, H. J., McKay, C. M., and Vandali, A. E. (**1992a**). ''A new portable sound processor for the University of Melbourne/Nucleus Limited Multichannel cochlear implant,'' J. Acoust. Soc. Am. **91**, 3367–3371.

McDermott, H. J., McKay, C. M., Vandali, A. E., and Clark, G. M. (**1992b**). ''A comparison of speech perception of cochlear implantees using the Spectral Maxima Sound Processor (SMSP) and the MSP (Multipeak) processor,'' Acta Otolaryngol. (Stockholm) **112**, 752–761.

McKay, C. M., O'Brien, A., and James, C. J. (**1999**). ''Effect of current level on electrode discrimination in electrical stimulation,'' Hear. Res. **136**, 159–164.

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

Nejime, Y., and Moore, B. C. J. (**1997**). ''Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise,'' J. Acoust. Soc. Am. **102**, 603–615.

Nelson, D. A., VanTassell, D. J., Schroeder, A. C., Soli, S., and Levine, S. (**1995**). ''Electrode ranking of 'place pitch' and speech recognition in electric hearing,'' J. Acoust. Soc. Am. **98**, 1987–1999.

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

Pfingst, B. E., Holloway, L. A., Zwolan, T. A., and Collins, L. M. (**1999**). ''Effects of stimulus level on electrode-place discrimination in human subjects with cochlear implants,'' Hear. Res. **134**, 105–115.

Robert, M. E. (**1998**). ''CONDOR: Documentation for Identification Test Program,'' (House Ear Institute, Los Angeles).

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,'' Hear. Res. **127**, 108–118.

Seligman, P., and McDermott, H. (**1995**). ''Architecture of the Spectra 22 Speech Processor,'' Ann. Otol. Rhinol. Laryngol. Suppl. **104**, 139–141.

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

Shannon, R. V., Jensvold, A., Padilla, M., Robert, M., and Wang, X. (**1999**). ''Consonant recordings for speech testing,'' J. Acoust. Soc. Am. (ARLO) **106**, L71–L74.

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

Skinner, M. W., Holden, L. K., Holden, T. A., Demorest, M. E., and Fourakis, M. S. (**1997**). ''Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants,'' J. Acoust. Soc. Am. **101**, 3766–3782.

Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (**1985**). ''Speech perception ability and psychophysical tuning curves in hearing impaired listeners,'' J. Acoust. Soc. Am. **77**, 620–627.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1992**). ''Effect of spectral envelope smearing on speech reception I,'' J. Acoust. Soc. Am. **91**, 2872–2880.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1993**). ''Effect of spectral envelope smearing on speech reception. II,'' J. Acoust. Soc. Am. **93**, 1547–1552.

Throckmorton, C. S., and Collins, L. M. (**1999**). ''Investigation of the effects of temporal and spacial interactions on speech-recognition skills in cochlear-implant subjects,'' J. Acoust. Soc. Am. **105**, 861–873.

Turner, C. W., and Henn, C. C. (**1989**). ''The relation between vowel recognition and measures of frequency resolution,'' J. Speech Hear. Res. **32**, 49–58.

Turner, C. W., Chi, S.-L., and Flock, S. (**1999**). ''Limiting spectral resolution in speech for listeners with sensorineural hearing loss,'' J. Speech Hear. Res. **42**, 773–784.

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 Hear. Res. **35**, 942–949.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). ''New levels of speech recognition with cochlear implants,'' Nature (London) **352**, 236–238.

Wilson, B. S., Finley, C. C., Lawson, D., and Zerbi, M. (**1998**). ''Temporal representations with cochlear implants,'' Am. J. Otol. **18** (Suppl.), S30–S34.

Zwolan, T. A., Collins, L. M., and Wakefield, G. H. (**1997**). ''Electrode Discrimination and speech recognition in postlingually deafened adult cochlear implant subjects,'' J. Acoust. Soc. Am. **102**, 3673–3685.

Zwolan, T. A., Kileny, P. R., Ashbaugh, C., and Telian, S. A. (**1996**). ''Patient performance with the Cochlear Corporation 20+2 implant: bipolar versus monopolar activation,'' Am. J. Otol. **17**, 717–723.