Effects of Presentation Level on Phoneme and Sentence Recognition in Quiet by Cochlear Implant Listeners

Gail S. Donaldson, and Shanna L. Allen

Objective: **The objectives of this study were to characterize the effects of presentation level on speech recognition in quiet by cochlear implant users with the Nucleus 22 SPEAK and Clarion v1.2 CIS speech-processing strategies, and to relate speech recognition at low presentation levels to stimulus audibility as measured by sound field thresholds. It was hypothesized that speech recognition performance in both Nucleus SPEAK and Clarion CIS participants would decrease as presentation level was decreased below 50 to 60 dBA, due to audibility limitations. However, it was expected that such level effects would be less severe in CIS participants than in SPEAK participants because the Clarion v1.2 device encodes a wider acoustic dynamic range (up to 60 dB) than the Nucleus 22 device (30 dB).**

Design: **Performance-intensity (P-I) functions for vowels, consonants and sentences in quiet were obtained from each participant. P-I functions incorporated speech levels of 70, 60, 50, 40 and 30 dBA. Subjects used their clinical speech processor maps and adjusted the loudness (volume/sensitivity) controls on their processors so that speech presented at 60 dBA was comfortably loud. Maps were created using default clinical procedures and were not adjusted to optimize sound field thresholds. Sound field thresholds and dynamic ranges were measured for warbled pure tones with frequencies of 250 to 6000 Hz.**

Results: **Consonant and sentence recognition showed strong level effects for both SPEAK and CIS participants, with performance decreasing substantially at levels below 50 dBA in most individuals. Vowel recognition showed weaker level effects. For all three speech materials, SPEAK and CIS participants demonstrated similar mean performance at 70 dBA; however, SPEAK participants showed larger reductions in performance than CIS participants with decreasing level. Sound field thresholds were more sensitive for CIS participants than for SPEAK participants, supporting the hypothesis that performance differences were related to audibility.**

Conclusions: **Cochlear implant listeners are unable to maintain good speech recognition at low presentation levels due to reduced stimulus audibility, and this may significantly limit their ability to communicate in daily life. It is likely that audibility differences between SPEAK and CIS participants in the present study can be attributed at least partly to differences in the acoustic dynamic range used by the respective processors. However, several additional factors may have contributed to differences in audibility and perception of soft speech among individual listeners with both devices. These include the minimum and maximum electrical stimulation levels specified in participants' maps and the speech processor sensitivity setting used for testing.**

(Ear & Hearing 2003;24;392–405)

Clinical Psychoacoustics Laboratory, Department of Otolaryngology, University of Minnesota, Minneapolis, Minnesota.

DOI: 10.1097/01.AUD.0000090340.09847.39

Because cochlear implant technology has improved in recent years, performance of the best cochlear implant users under ideal listening conditions has begun to approach that of normal-hearing listeners (Skinner et al., 1994; Wilson, Reference Note 1). However, when listening conditions are less favorable, implant users may encounter much greater difficulty than their normal-hearing counterparts. For example, even small amounts of background noise can cause implant listeners to experience substantial reductions in speech recognition (Fu, Shannon, & Wang, 1998; Hochberg, Boothroyd, Weiss, & Hellman, 1992).

Relatively little is known about cochlear implant listeners' ability to perceive speech at levels lower than those typically evaluated clinically (50 to 70 dB SPL). An impressive feature of normal acoustic hearing is its wide dynamic range, which permits encoding of a 1,000,000-fold (120 dB) range of sound pressures and allows normal-hearing listeners to maintain excellent speech recognition over a wide range of speech presentation levels. Because cochlear implant speech processors encode only a limited portion of the acoustic dynamic range (typically 30 to 60 dB), implant listeners have access to a narrower range of acoustic sound intensities than normal-hearing listeners. Low-level sounds can be emphasized by adjusting the sensitivity control on

0196/0202/03/2405-0392/0 • Ear & Hearing • Copyright © 2003 by Lippincott Williams & Wilkins • Printed in the U.S.A.

the speech processor; however, it is not feasible for the implant listener to adjust this control continuously during the course of a conversation, for example, to follow the speech of two speakers at different distances. Instead, the listener typically adjusts the sensitivity control to an appropriate setting for moderately loud speech, which limits his ability to perceive low-level speech cues and soft speech.

Skinner, Holden, Holden, Demorest, and Fourakis (1997) evaluated cochlear implant listeners' speech recognition in quiet for presentation levels corresponding to soft, conversational, and raised-to-loud speech (50, 60, and 70 dB SPL, respectively). Strong level dependencies were observed for consonants, words, and sentences: Scores decreased systematically as presentation level was reduced from 70 to 50 dB SPL, and scores at 50 dB were often markedly lower than those at the two higher levels. Measurement of sound field thresholds indicated that many acoustic cues were perceived by listeners as being soft or inaudible at 50 dB SPL, limiting speech recognition at that level. Vowel recognition showed a weaker dependence on presentation level, presumably because vowel cues possess more energy than consonant cues and are thereby less susceptible to audibility effects. Maximum vowel recognition scores were typically obtained at 60 dB SPL, with slightly lower scores obtained at both 50 and 70 dB SPL. Reasons for the decrease in performance that occurred between 60 and 70 dB SPL were unclear, but the authors speculated that it could be attributable to compression of the highest speech peaks by the processor's automatic gain control. The strong level dependence observed for consonant, word, and sentence recognition led the authors to recommend that clinical testing be performed at 60 dB SPL rather than 70 dB SPL, because 60 dB SPL is more typical of speech levels encountered in everyday life. The data of Skinner et al. were obtained in listeners with the Nucleus 22 cochlear implant who used the SPEAK speech processing strategy. Similar findings were reported by Muller-Deile, Schmidt, and Rudert (1995) in Nucleus 22 listeners using the MPEAK strategy. To our knowledge, the effects of presentation level on speech recognition have not been reported for cochlear implant listeners with other devices, nor have presentation levels lower than 50 dB SPL been evaluated. Because the Nucleus 22 implant uses a relatively narrow acoustic dynamic range, the results of Skinner et al. and Muller-Deile et al. may not generalize to implant users with other devices; thus, it is important to assess presentation-level effects for users of other cochlear implants. It is also important to assess speech recognition for presentation levels lower than 50 dB SPL, because low-level speech is encountered routinely in daily life.

Several factors are likely to influence cochlear implant listeners' ability to hear and recognize soft speech. These include the size of the acoustic dynamic range encoded by the speech processor, the setting selected for the sensitivity control on the speech processor, the shape of the acoustic-to-electric amplitude-mapping function, and the procedure used to set electrical thresholds and maximum stimulation levels.

The effects of acoustic dynamic range or "input dynamic range" (IDR) have been evaluated in two recent studies. Fu and Shannon (1999) measured vowel and consonant recognition in three Nucleus-22 listeners using a 4-channel CIS processor. Mean performance in quiet decreased as the input dynamic range was reduced from 40 dB to values less than 30 dB, either by peak clipping or center clipping. Average reductions in performance became statistically significant when IDRs were 20 dB or lower for vowels, and 14 dB or lower for consonants. These results led the authors to conclude that a 30-dB IDR provides adequate speech information for phoneme identification, even when spectral resolution is restricted to four spectral channels. More recently, Zeng et al. (2002) measured the effects of IDR on vowel and consonant recognition in Clarion cochlear implant users with the CIS strategy, for IDR settings ranging from 10 to 60 dB. In contrast to the Fu and Shannon results, these authors found that IDRs of 50 to 60 dB yielded the highest scores for consonant recognition in quiet, and that IDRs of 40 to 50 dB yielded the highest scores for vowel recognition in quiet. Findings were consistent with measures of speech dynamic range obtained in the same study: Envelope amplitudes in narrow frequency bands were found to have dynamic ranges of 41 to 52 dB across channels for consonants and dynamic ranges of 36 to 53 dB across channels for vowels. Because this study used similar speech materials as the Fu and Shannon study (multitalker consonants and vowels), it is not clear why the two studies obtained such different findings. Unpublished data obtained in our own laboratory are consistent with Zeng et al.'s findings that IDRs of 40 to 60 dB are optimal for speech recognition in quiet (Allen & Donaldson, Reference Note 2).

With respect to amplitude mapping, Fu and Shannon (1998) investigated the effects of varying the exponent of a power-law mapping function on vowel and consonant recognition in cochlear implant users with a 4-channel CIS processor and in normal-hearing listeners with a 4-channel CIS simulation. In this study, input dynamic range was fixed at 40 dB. Exponents of the power-law mapping function were varied from 0.05 to 0.75 for cochlear implant users and from 0.3 to 3 for normal-hearing listeners. For the cochlear-implant listeners, an exponent of 0.2 yielded the best performance; for normal-hearing listeners, a power-function exponent of 1 (linear mapping) was optimal. However, for both types of listeners, the use of exponents larger or smaller than the optimal ones yielded only slightly poorer performance. This led the authors to conclude that the compression exponent of the amplitude-mapping function has only a small effect on phoneme recognition in quiet. Similar results were reported by Zeng and Galvin (1999), who showed that amplitude compression had no systematic effect on vowel or consonant recognition in cochlear implant users with the SPEAK strategy. Both of the above studies stimulated cochlear implant listeners at comfortably loud levels, thus, it is not clear whether their findings would generalize to softer speech. It seems possible that greater compression may be beneficial in the case of soft speech, because increasing the compression exponent causes low-level acoustic information to be mapped to higher current levels within the electric dynamic range, thereby improving audibility.

Skinner and her colleagues (Skinner, Holden, Holden, & Demorest, 1995; Sun, Skinner, Liu, Wang, Huang, & Lin, 1998; Skinner, Holden, Holden, & Demorest, 1999) have shown that the specific procedures used to set minimum and maximum electrical stimulation levels can also affect implant listeners' ability to hear soft speech. Skinner et al. (1999) compared the use of thresholds versus raised levels of minimum stimulation in eight Nucleus 22 users with the SPEAK strategy. Word recognition in quiet and sentence recognition in noise were measured for speech presentation levels of 50, 60 and 70 dB SPL. Use of a raised level program resulted in small but significant increases in speech recognition for words presented at 50 and 60 dB SPL and for sentences presented at 50 and 70 dB SPL. Loudness judgments obtained with the two programs indicated that the use of raised levels increased the loudness of low-level speech sounds, thereby making them more salient to the listener. These findings indicate that minimum electrical stimulation levels should be set above regions of slow loudness growth that may exist near threshold. In other work, this group has demonstrated the importance of optimizing maximum electrical stimulation levels to make conversational speech comfortably loud when the speech processor is set at a moderate sensitivity control setting, and to avoid distortion of loud sounds (Skinner et al., 1995; Sun et al., 1998; Seligman & Whitford, 1995). As discussed later, the ability to make optimal use of higher sensitivity settings to perceive soft speech depends on the use of appropriate maximum stimulation levels. Thus, the settings for minimum as well as maximum stimulation levels may have a significant effect on cochlear implant listeners' ability to recognize low-level speech.

In undertaking the present study, we had three major goals: The first goal was to evaluate the effects of presentation level on speech recognition in quiet by cochlear implant listeners, extending presentation levels to lower values than those studied previously by Skinner, Holden, Holden, Demorest, and Fourakis (1997). The second goal was to compare presentation-level effects for Nucleus 22 listeners with those for Clarion v1.2 listeners. Finally, the third goal was to relate level effects to audibility as measured by sound field thresholds. Our general hypothesis was that audibility limitations and corresponding level effects in speech recognition would be greater for Nucleus listeners than for Clarion listeners as a consequence of the narrower acoustic dynamic range implemented in the Nucleus device.

METHODS

Participants

Participants in the study were 14 postlingually deafened adults with cochlear implants. Seven used a Nucleus 22 cochlear implant with the SPEAK speech processing strategy (SPEAK group), and 7 used a Clarion v1.2 cochlear implant with the continuous interleaved sampler (CIS) speech processing strategy (CIS group). Each participant exhibited at least moderate open-set speech recognition in quiet with their device. Table 1 provides relevant demographic information, including gender, age at the time of testing, etiology of deafness, duration of deafness before implantation, and duration of cochlear implant use. Average duration of deafness was longer for CIS users than for SPEAK users (23.7 yr versus 3.1 yr), and CIS users had fewer years of experience with their devices (1.5 yr versus 5.4 yr). Although these factors predict poorer speech recognition for the CIS group (Blamey et al., 1996), the two groups demonstrated similar average performance at the highest presentation levels tested (Figs. 1, 3, 5).

Phoneme P-I Functions

Performance-intensity (P-I) functions for vowel and consonant recognition were obtained using a standard phoneme-confusion procedure. Vowel stimuli were 11 /h-V-d/ monosyllables from the database of Hillenbrand, Getty, Clark, and Wheeler

TABLE 1. Participant code, gender, age, etiology of hearing loss (HL), duration of severe-to-profound bilateral HL before implantation, years cochlear implant (CI) use, number of channels programmed in speech processor map, sensitivity and volume settings used during testing, speech processor input dynamic range, frequency range of speech processor map, mean electrical threshold and dynamic range for active electrodes (from the speech processor map), and mean sound field thresholds and dynamic range (500–4k Hz)

Code	M/F	Age	Etiology	Yrs Deaf	Yrs CI use	# chnls	Sens	Vol	IDR (dB)	Freq range (Hz)	Mn elec THS (Clin Units)	Mn elec DR (Clin Units)	Mn SF THS (dB SPL)	Mn SF DR (dB)
Nucleus 22 SPEAK participants														
N09	M	56	Meniere's disease	1	10.1	19	3	—	30	120–7390	101.5	47.5	47.6	32.9
N12	M	41	progressive SNHL	8	8.4	17	5.5	—	30	150–6730	107.1	42.4	34.1	22.9
N13	M	52	progressive SNHL	4	8.7	18	3.5	—	30	120–6308	132.9	34.7	36.2	30.0
N14	M	49	progressive SNHL	1	4.9	20	5	—	26	120–8658	86.4	55.0	41.2	22.1
N28	M	57	meningitis	< 1	2.5	20	1.5	—	30	120–8658	89.6	54.6	51.2	12.1
N30	F	57	otosclerosis	10	1.6	11	1.5	—	30	240–4288	66.4	79.5	46.2	21.4
N32	M	29	maternal rubella	< 1	1.6	18	2	—	30	150–7885	58.7	73.0	41.2	25.7
Clarion 1.2 CIS participants														
C01	F	42	progressive SNHL	13	2.4	8	10:30	12:00	50	250–5500	134.3	343.1	26.9	31.4
C02	M	37	unknown	19	1.6	8	10:30	12:00	58	250–6800	60.0	293.8	18.4	37.9
C03	F	49	hereditary SNHL	27	0.9	8	10:30	12:30	60	350–5500	111.8	259.6	23.4	39.3
C05	M	42	unknown	< 1	0.9	8	10:30	11:30	50	250–5500	121.9	187.1	27.6	32.9
C07	F	58	hereditary SNHL	35	0.9	8	10:30	12:00	50	350–5500	86.7	111.9	23.4	31.4
C08	F	31	ototoxicity	26	3.3	8	10:30	1:00	60	250–5500	97.4	213.0	34.8	30.0
C09	M	41	unknown	41	0.5	8	10:30	12:00	60	350–5500	97.4	71.8	38.4	32.1

Clinical units are SUs (stimulus units) for Nucleus participants and CUs (clinical units) for Clarion participants.

(1995). Stimuli spoken by 6 different male talkers were selected for each vowel, with the entire stimulus set representing 20 male talkers*. Vowels tested were /æ, ɑ, ɛ, e, ɝ, I, i, o, ʊ, ʌ, u/as in "*had*, *hod*, *head*, *hayed*, *heard*, *hid*, *heed*, *hoed*, *hood*, *hud* and *who'd*." Consonant stimuli were 19 /a-C-a/ disyllables from the stimulus set of Van Tasell, Greenfield, Logemann, and Nelson (1992), spoken by three male and three female talkers. Consonants tested were /p, t, k, b, d, g, f, θ, s, ʃ, v, ð, z, ʒ, m, n, r, l, j/. Sampling rates for the vowel and consonant stimuli were 16 kHz and 10 kHz, respectively, and both were digitized with 12-bits of amplitude resolution.

Phoneme-confusion testing was conducted in a sound-isolated room, with the listener seated approximately 1 meter in front of a pair of high-quality loudspeakers and a video screen. Digitized speech tokens were played out from computer memory, amplified, low-pass filtered at half the digitization rate, attenuated, and presented through the speakers. The stimulus was presented once on each trial, and the listener used a computer mouse to select his or her response from a list of possible alternatives displayed on the video screen. Correct-answer feedback was provided immediately after each response.

Vowel and consonant data were obtained in separate test sessions. In each case, three sets of data were obtained. Each set consisted of one block of phonemes at each presentation level (70, 60, 50, 40, and 30 dBA), presented in descending order. In turn, each block was comprised of 6 trials per phoneme (66 vowels or 114 consonants) presented in random sequence. Thus, a total of 18 trials were completed for each phoneme at each presentation level.

Merged confusion matrices were analyzed using information transmission analysis (Miller & Nicely, 1955) to obtain transmitted information (TI) measures for overall consonant and vowel recognition, and for specific vowel and consonant features. Feature categories used for these vowel and consonant analyses are shown in Tables 2 and 3, respectively. Vowel categories were height (related to f1 frequency), frontness (related to f2 frequency) and duration. Category membership was established using values of f1, f2, and duration measured by Hillenbrand et al. for the specific tokens used in this study. Based on the distribution of measured values, three categories were used for vowel height, and two categories each were used for vowel frontness and duration. The ranges of f1 values measured across tokens were 312 to 453 Hz (low f1), 424 to 559 Hz (mid f1), and 594 to 825 Hz (high f1); the ranges of f2 values were 868 to 1377 Hz (low f2), and 1851 to 2311 Hz (high f2); and the ranges of vowel durations were 156 to 241 msec (short) and 218 to 342 msec (long). With one token eliminated from each of

*Only male talkers were used for vowel testing because the female talkers in the Hillenbrand data set exhibit wide, sometimes overlapping, ranges of formant frequencies for a given vowel. Although this variability can occur for both male and female talkers, it was much less pronounced for the male talkers used here.

the short and long duration categories, the range of vowel durations was non-overlapping (156 to 201 msec for short; 218 to 342 msec for long), so that vowel duration was a viable cue to vowel identity. Consonant categories were voicing, manner and place-of-articulation. Standard category membership was used.[†]

Sentence P-I Functions

P-I functions were also obtained for HINT sentences in quiet (Nilson, Soli, & Sullivan, 1994). Two 10-sentence lists were obtained at each of the five presentation levels (70, 60, 50, 40 and 30 dBA), with the order of levels randomized independently for each listener. Sentence stimuli were routed from an audio compact disc player through an audiometer to a single high-quality loudspeaker placed 1 meter in front of the listener's chair. Listeners repeated each sentence verbally and the tester recorded the number of words repeated correctly on a printed score sheet. The number of words repeated correctly for the 20 sentences presented at a given stimulus level were converted to a percent-correct score.

Level Calibration

Consonant and vowel stimuli were calibrated by measuring speech peaks on the slow A scale of a Bruel and Kjaer Type 2203 sound level meter.[‡] An attenuation value was determined that resulted in an average peak level of 60 dBA, and attenuation was altered in 10 dB steps to produce average peak levels of 70, 50, 40 and 30 dBA. HINT sentences were calibrated using the 1-kHz calibration tone provided on the audio compact disc.

[†]The place feature for /j/ was categorized as "mid," as in Van Tasell et al. (1992); it is sometimes categorized as "back."

[‡]Repeated measurements of slow-peak levels for a given stimulus typically yielded values that differed by no more than 1 dB. An alternative approach used by many investigators is to equate stimuli on the basis of RMS amplitude. RMS measures of our peak-equated stimuli (excluding silent segments that preceded or followed the stimulus itself in the digitized waveform file) varied over an 8-dB range for consonants and a 10-dB range for vowels.

TABLE 2. Vowel feature categories used for information transmission analyses in the present study

	Height (f1 frequency)	Frontness (f2 frequency)	Duration
1	/ I, i, u / (high)	/ ɑ, ɚ, o, ʊ, ʌ, u / (low)	/ ɛ, I, ʊ, ʌ / (short)
2	/ ɛ, e, ɚ, o, ʊ / (mid)	/ æ, ɛ, e, I, i / (high)	/ æ, ɑ, e, ɚ, i, o, u / (long)
3	/ æ, ɑ, ʌ / (low)		

TABLE 3. Consonant feature categories used for information transmission analyses in the present study

	Voicing	Manner	Place
1	/ b, d, g, v, ð, z, ʒ, m, n, r, l, j / (voiced)	/ p, t, k, b, d, g / (stops)	/ b, p, f, v, m / (front)
2	/ p, t, k, f, θ, s, ʃ / (voiceless)	/ f, θ, s, ʃ, v, ð, z, ʒ / (fricatives)	/ t, d, θ, s, ð, z, n, l, j / (mid)
3		/ m, n / (nasals)	/ k, g, ʃ, ʒ, r / (back)
4		/ r, l, j / (glide-liquids)	

Speech Processing Strategies

As indicated earlier, Nucleus participants used the spectral peak (SPEAK) strategy implemented on a Spectra speech processor. The SPEAK strategy divides the incoming acoustic signal into a maximum of 20 frequency bands and samples the envelope amplitude in each band approximately once every 4 msec. The 20 frequency bands are assigned to 20 bipolar electrode pairs ordered tonotopically along the internal array. On each stimulation cycle, current pulses are delivered to those electrodes corresponding to the filter bands with the highest envelope amplitudes. The average number of electrodes stimulated per cycle is specified in the speech processor program. The default value of 6 was used by all Nucleus participants in the present study except N30, who used a value of 5 because she had a reduced number of stimulable electrodes. The SPEAK strategy has a pulse rate of approximately 250 pulses/sec per electrode, corresponding to the 4-msec stimulation cycle. Additional details concerning this strategy may be found in Skinner et al. (1994). Clarion participants used the continuous interleaved sampler (CIS) processing strategy implemented on an S-series processor. This strategy divides the incoming acoustic signal into 8 frequency bands that are assigned to 8 monopolar-coupled electrodes along the internal array. Unlike the SPEAK strategy, the CIS strategy stimulates every electrode on every stimulation cycle, regardless of the envelope amplitude in a given channel. Pulse rate is approximately 800 pulses/sec per electrode. A detailed description of the CIS strategy is given by Wilson, Finley, Lawson, Wolford, Eddington, and Rabinowitz (1991). The SPEAK and CIS strategies are designed to emphasize spectral and temporal aspects of the speech signal, respectively. We are not aware of any evidence that the relative importance of spectral and temporal speech cues varies with speech presentation level; thus, it is not clear that one strategy would have an inherent advantage over the other for the perception of soft speech.

Speech Processor Settings

Participants used their own speech processors and clinical speech processor maps for all testing. Information concerning the number of frequency channels activated in each participant's map, the range of frequencies encoded by the processor, and the mean electrical threshold and dynamic range for activated electrodes is included in Table 1. Clinical maps were created using standard fitting software and our clinic's standard mapping procedures. Electrical thresholds (T-levels) were measured using an ascending method of adjustment and were specified as the current level near threshold at which loudness began to increase reliably with increasing current amplitude (i.e., T-tails were eliminated). Maximum acceptable comfort levels (C-levels, Nucleus device) and comfortable loudness levels (M-levels, Clarion device) were also measured with an ascending method of adjustment procedure. C or M levels were matched in loudness by balancing the loudness of electrodes, two at a time. For Nucleus users, loudness was also balanced at a level corresponding to 90% of the electrical dynamic range, 5 electrodes at a time, and necessary adjustments were made by increasing or decreasing T levels. Loudness was not balanced for T-levels. Once loudness-balancing procedures were completed, the participant was stimulated with speech. If speech was too loud, C or M levels were reduced globally across all electrodes. If speech was too soft, T levels were increased globally (Nucleus device) or M levels were increased globally (Clarion device).

The Nucleus Spectra speech processor has a single control that can be adjusted by the listener to vary loudness. This is the sensitivity control, which alters the gain applied to the input signal. For this study, Nucleus participants adjusted the sensitivity control on their Spectra speech processor to a setting that produced comfortable loudness for a 60-dBA speech input. This sensitivity setting was then held constant for all speech materials and presentation levels. The Clarion v1.2 speech processor has both a sensitivity control and a volume control. The sensitivity control alters input gain and determines the acoustic amplitude at which AGC compression is invoked. In the present study, the sensitivity control was fixed at "10:30" to eliminate AGC compression at stimulus levels less than approximately 70 dBA. The volume control on the Clarion device varies the upper limit of current amplitudes presented to the electrodes. Clarion participants adjusted the volume control to a setting that produced comfortable loudness for a 60-dBA speech input. This volume setting was then held constant for all speech materials and presentation levels. In most cases, the sensitivity and volume settings used by individual participants were similar to those used in daily life.

As indicated previously, one speech-processor characteristic having special relevance to the present study is the input dynamic range (IDR), or range of speech-envelope amplitudes (in dB) that is mapped into the electrical dynamic range of participants' implanted electrodes. The Clarion 1.2 device allows the IDR to be set to values between 20 and 60 dB. The default setting is 60 dB; however, some listeners use a reduced setting to eliminate the perception of low-level device noise. Clarion participants in the present study used IDR settings ranging from 50 to 60 dB (Table 1). The Nucleus 22 device controls the acoustic dynamic range indirectly through the base level parameter, with the maximum acoustic dynamic range limited to approximately 30 dB. With one exception, Nucleus participants in the present study used a base level setting of 4, corresponding to an acoustic dynamic range of 29.5 dB. The remaining participant (N14), used a base level setting of 7, resulting in an acoustic dynamic range of 25.5 dB.

Sound Field Measures

Sound field thresholds were measured for warbled pure tones at half-octave frequencies between 250 Hz and 6k Hz, using a standard audiometric procedure (Carhart and Jerger, 1959) with a 5 dB step size. "Medium loud" levels (MLLs) were measured at the same frequencies, by increasing stimulus level in 5 dB steps until listeners' loudness ratings exceeded "medium loud" and reached "loud." The highest stimulus level rated as "medium loud" at each frequency was taken to be the MLL.

RESULTS

Vowel P-I Functions

Figure 1 shows performance-intensity (P-I) functions for vowel recognition, with performance expressed as percent transmitted information (% TI). Data for individual SPEAK and CIS listeners are plotted in the left and center panels, respectively, and mean data for both groups are plotted in the right panel. The right panel also shows mean data for two normal-hearing acoustic listeners who completed the same protocol as the cochlear implant participants. Note that the acoustic listeners achieved nearly perfect vowel-recognition scores at all presentation levels.

SPEAK participants achieved moderate to high vowel recognition scores, ranging from 60% to 96%TI, at presentation levels of 50 to 70 dBA, but showed dramatic decreases in performance at lower

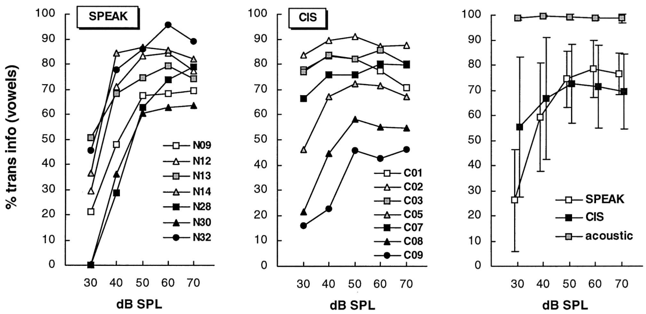

Figure 1. Performance-intensity functions for vowel recognition, expressed as percent transmitted information. Individual data for SPEAK and CIS listeners are shown in the left and center panels, respectively. Mean data for SPEAK and CIS listeners, together with mean data for two normal-hearing listeners, are shown in the right panel. Error bars indicate ± 1 SD.

levels. The four best-performing participants (N12, N13, N14, and N32) maintained relatively high levels of performance at 40 dBA, but exhibited sharp decreases in performance as presentation level was reduced to 30 dBA. The three remaining participants showed clear decrements in performance even at 40 dBA. All SPEAK participants reported that the 30 dBA stimuli were very soft, and two of them (N28 and N30) reported that most or all of the vowel stimuli at this level were inaudible.

CIS participants showed a wider range of performance than SPEAK participants, with vowel recognition ranging from 43% to 91% TI at presentation levels of 50 to 70 dBA. Presentation level also had a more variable effect on performance in this group. The four best-performing participants (C01, C02, C03 and C07) maintained high levels of performance at all stimulus levels, the participant with the next highest performance (C05) showed a decrement in performance only at 30 dBA, and the two poorest-performing participants (C08, C09) showed decrements at 30 and 40 dBA. The mean group data showed nearly constant performance for presentation levels of 40 to 70 dBA and showed only a small decrement at 30 dBA.

Note that P-I functions were non-monotonic for some cochlear implant listeners (e.g., N12, C01), with optimum performance achieved at stimulus levels of 60, 50 or even 40 dBA, and decreasing at higher levels. As discussed earlier, this "rollover" effect was also observed by Skinner, Holden, Holden, Demorest, and Fourakis (1997) and may reflect a loss of spectral resolution due to compression of speech peaks at the highest stimulus levels.

A 2-way (group x level) repeated-measures ANOVA applied to the vowel TI scores showed a significant main effect of presentation level ($F[4,48]=46.96$, $p < 0.001$) and a significant group x level interaction ($F[4,48]=13.2$, $p < 0.001$). For SPEAK listeners, vowel scores at 30 dBA were significantly poorer than those at 40 dBA, and scores at 40 dBA were significantly poorer than scores at 50, 60 and 70 dBA (Tukey tests, $p < 0.001$). For CIS listeners, vowel scores at 30 dBA were significantly poorer than those at 50, 60 and 70 dBA (Tukey tests, $p < 0.05$). Mean vowel scores for SPEAK and CIS listeners were significantly different at only one presentation level, 30 dBA (Tukey test, $p < 0.05$).

Figure 2 shows group transmitted information scores for the F1, F2 and duration (DUR) vowel features. The strongest effect of presentation level was observed for the F2 feature in SPEAK listeners, with TI scores decreasing systematically from a very high value of 90.2% at 70 dBA to a low value of 17.3% at 30 dBA. The SPEAK functions for the F1 and DUR features showed a weaker level dependence than the F2 feature but, as expected, showed greater decrements than the corresponding CIS functions at low stimulus levels (30 to 40 dB). Notably, F1 information was transmitted more poorly than F2 information for both SPEAK and CIS listeners. This was true not only for the mean data shown in Fig. 2 but also for every individual listener in each group. F1 transmission may be poorer than F2 transmission because F1 spans a narrower frequency range (327 to 809 Hz for the present stimuli) than F2 (891 to 2131 Hz) and is less well resolved by the speech encoding strategies. Skinner, Holden, and Holden (1997) showed that increasing from four to six the number of electrodes assigned to the F1 frequency region improved F1 transmission with the SPEAK strategy. SPEAK maps used by participants in the present study assigned 4 frequency channels to F1 frequencies and 7 channels to F2 frequencies, whereas CIS maps assigned 3 frequency channels to F1 frequencies and 4 frequency channels to F2 frequencies. Transmission of F1 information was especially poor in two CIS participants in this study (C08 and C09), never exceeding 25% TI at any

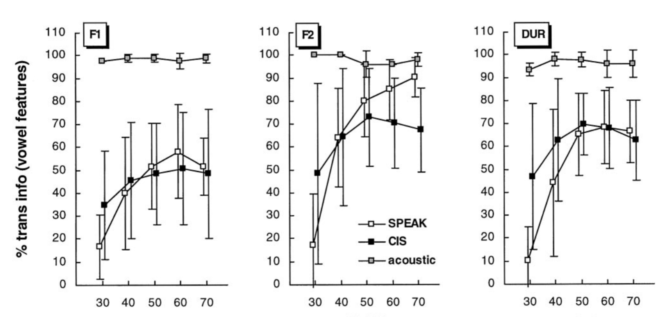

Figure 2. Performance-intensity functions for the vowel features F1, F2 and duration (left, middle and right panels, respectively). Each panel shows mean data for SPEAK, CIS and normal-hearing listeners. Error bars indicate ± 1 SD.

presentation level. In these listeners, unfavorable patterns of neural survival may have reduced the functional number of frequency channels coding F1 information below the three nominal channels provided by the processor.

Consonant P-I Functions

Figure 3 shows P-I functions for consonant recognition. Similar to previous figures, the data for SPEAK and CIS listeners are plotted in the left and center panels, and mean data for both cochlear implant groups and for two normal-hearing listeners are plotted in the right panel. Note that the normal-hearing listeners again achieved high levels of performance at all stimulus levels, although their overall performance was slightly lower for consonants (91%) than for vowels (99%). Individual SPEAK listeners obtained moderate to high consonant recognition scores, ranging from 59 to 83%TI, at the highest presentation level of 70 dBA. For each listener, performance remained constant or decreased slightly at 60 dBA, but then decreased dramatically at 50 dBA and lower presentation levels. As was the case for vowel stimuli, the 30 dBA consonant stimuli were described as being very soft or inaudible, and only one of the seven SPEAK listeners achieved a non-zero score at the 30 dBA level. CIS listeners again showed a wider range of performance than SPEAK listeners but showed a weaker effect of presentation level. Performance was relatively constant at the two highest levels of 60 and 70 dBA, ranging from 46 to 83%TI, but decreased gradually as presentation level was reduced to 50 dBA and lower levels. The shapes of individual P-I functions were more consistent across listeners for consonants than was the case for vowels.

A 2-way (group x level) repeated-measures ANOVA applied to the consonant TI data confirmed a significant main effect of presentation level ($F[4.40]=181.6$, $p < 0.001$) and a significant group x level interaction ($F[4,40]=20$, $p < 0.001$). For SPEAK participants, consonant %TI scores were similar at 60 and 70 dBA, but performance at each of the lower presentation levels (50, 40 and 30 dBA) was significantly poorer than that at the next highest level (Tukey tests, $p < 0.001$). For CIS participants, performance was constant at 50, 60 and 70 dBA, but performance at 30 dBA was significantly poorer than that at 40 dBA and higher levels, and performance at 40 dBA was significantly poorer than performance at 50, 60 and 70 dBA (Tukey tests, $p < 0.05$). Scores for CIS participants were significantly higher than those for SPEAK participants at the two lowest presentation levels, 30 and 40 dBA (Tukey tests, $p < 0.05$).

Group transmitted information scores for the consonant features of voicing, manner and place are shown in Figure 4. Both SPEAK and CIS listeners showed strong presentation-level effects for all three features. Mean performance was similar for the two groups at the highest presentation levels, but SPEAK participants performed more poorly than CIS participants at lower presentation levels. Of the three features, the highest scores were obtained for manner, with SPEAK and CIS listeners obtaining mean scores of 71.9 and 79%TI, respectively, at 70 dBA. Performance for the manner feature increased monotonically with level in both participant groups, with no clear evidence of saturation at the highest level. This was true both for the mean data (Fig. 2) and for most individual listeners. Exceptions occurred for three CIS participants who achieved asymptotic performance on the manner feature at 50 to 60 dBA. Slightly lower scores were obtained for the voicing feature, with SPEAK and CIS participants obtaining scores of 66.7 and 68.7% TI at 70

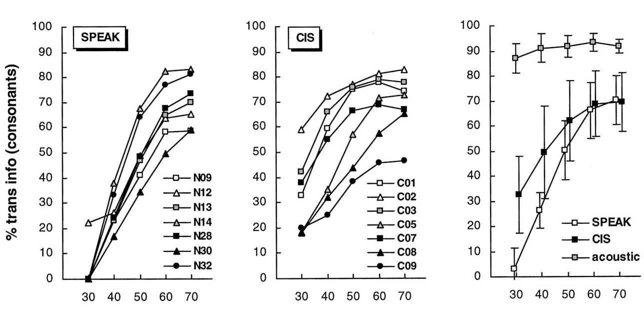

Figure 3. Performance-intensity functions for consonant recognition, expressed as percent transmitted information. Individual data for SPEAK and CIS listeners are shown in the left and center panels, respectively. Mean data for SPEAK and CIS listeners, together with mean data for two normal-hearing listeners, are shown in the right panel. Error bars indicate ± 1 SD.

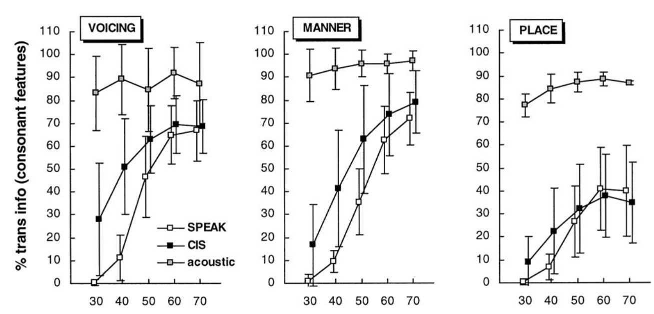

Figure 4. Performance-intensity functions for the consonant features voicing, manner and place-of-articulation (left, middle and right panels, respectively). Each panel shows mean data for SPEAK and CIS listeners, together with mean data for two normal-hearing listeners. Error bars indicate ± 1 SD.

dBA. Unlike the manner feature, mean performance on the voicing feature asymptoted at presentation levels of 50 to 60 dBA. Performance on the place feature was considerably poorer than performance on the manner and voicing features, with SPEAK and CIS listeners achieving average scores of 40.1 and 34.8% TI, respectively, at 70 dBA. As was the case for voicing, mean P-I functions for the place feature asymptoted at levels of 50 to 60 dBA, suggesting that factors other than audibility limited performance on this feature at the highest presentation levels. Again, the individual data were relatively homogeneous: Five of seven SPEAK listeners and all seven CIS listeners reached asymptotic performance on the place feature at presentation levels of 60 dBA or lower. The two remaining SPEAK listeners showed continued improvement in this feature as level was raised to 70 dBA. The two normal-hearing listeners showed relatively high, although not perfect, scores on all three features. The most common errors among the normal-hearing listeners were place-of-articulation errors for voiceless fricatives. There was no apparent level-effect in the normal-hearing data for the voicing and manner features, but a small level effect was observed for the place feature.

Sentence P-I Functions

Figure 5 shows the individual and mean P-I functions for HINT sentences. Sentence performance reflected performance for vowel and consonant recognition (Figs. 1 and 3) in that the listeners within each group ordered themselves similarly across the three tasks. Consistent with the vowel and consonant data shown in earlier figures, the average performance of CIS and SPEAK listeners was equivalent at 70 dBA, but CIS listeners maintained higher levels of performance than SPEAK listeners at lower presentation levels. Performance differences were particularly striking at presentation levels of 40 and 50 dBA, where mean CIS scores exceeded mean SPEAK scores by approximately 30%. Sentence P-I functions were more similar in shape to the consonant P-I functions (Fig. 3) than to the vowel P-I functions (Fig. 1), suggesting that consonant audibility was an important factor related to the recognition of words in sentences. The relationship between audibility and speech recognition is well established in acoustic hearing (French and Steinberg, 1947).

A 2-way (group x level) repeated-measures ANOVA applied to the arcsine-transformed sentence percent-correct data confirmed a significant main effect of presentation level ($F[4,48]=49.1$, $p < 0.001$). The group x level interaction failed to reach significance ($F[4,48]=2.5$, $p = 0.054$). Mean scores were not significantly different for presentation levels of 70 dBA versus 60 dBA or for presentation levels of 60 dBA versus 50 dBA. However, for all other pairs of presentation levels, mean scores were significantly greater for the higher presentation level as compared to the lower presentation level (Tukey test, $p < 0.05$).

Sound Field Measures

Figure 6 shows individual and mean sound field thresholds for Nucleus SPEAK and Clarion CIS participants, following the same format used in earlier figures. Individual sound field thresholds varied over a range of approximately 20 dB for participants within each group. However, for the primary speech frequencies (500 Hz to 4kHz), average thresholds for CIS participants (27.5 dB SPL) were 15 dB lower than those for SPEAK participants (42.5 dB SPL). The sharp rise in CIS sound field thresholds at 250 Hz reflects use of a more restricted low-frequency limit in Clarion maps (250 or 350 Hz) as compared to Nucleus maps (120 to 240 Hz) (see Table 1). It should be noted that average sound field

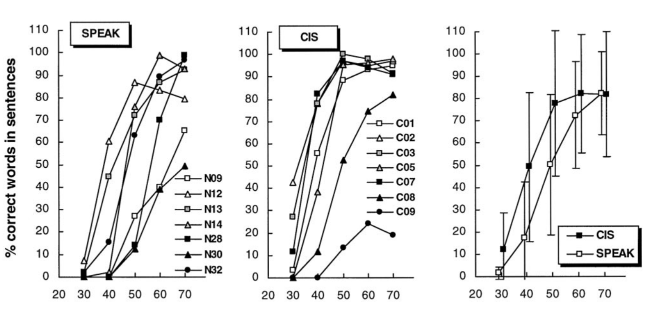

Figure 5. Performance-intensity functions for HINT sentences, expressed as percent correct words in sentences. Individual data for SPEAK and CIS listeners are shown in the left and center panels and mean data for both groups are shown in the right panel. Error bars indicate ± 1 SD.

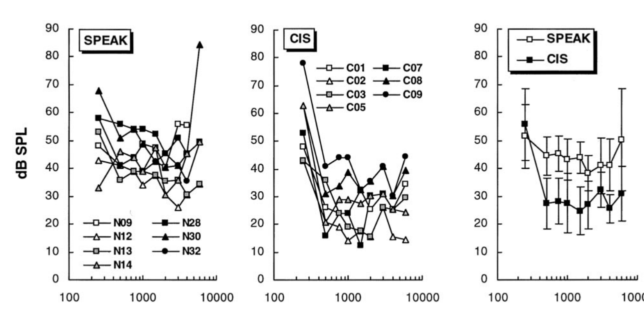

Figure 6. Sound field thresholds for warbled pure tones. Individual data for SPEAK and CIS listeners are shown in the left and center panels and mean data for both groups are shown in the right panel. Error bars indicate ± 1 SD.

thresholds for our SPEAK listeners were approximately 5 dB higher (less sensitive) than those observed by Skinner, Holden, Holden, Demorest, and Fourakis (1997). This difference may reflect variability in the sensitivity settings preferred by individual listeners as well as differences in procedures used to set electrical thresholds (T-levels) in listeners' maps.

The sound field thresholds shown in Figure 6 suggest that decrements in speech recognition at low presentation levels can be attributed largely to audibility limitations. SPEAK listeners show average speech-frequency sound field thresholds of 42.5 dB SPL. This is consistent with sharp decrements in vowel recognition at and below 40 dBA (Fig. 1) and sharp decrements in consonant recognition at and below 50 dBA (Fig. 3). CIS listeners have more sensitive speech-frequency sound field thresholds (average 27.5 dB SPL) and maintain better speech recognition scores at lower presentation levels. Audibility also accounts for some of the individual differences in level effect observed. Among SPEAK listeners, N28 and N30 had the highest soundfield thresholds and showed the sharpest reductions in vowel and consonant performance with decreasing stimulus level (Figs. 1, 3). Similarly, CIS listeners C08 and C09 exhibited the highest sound field thresholds and showed the largest overall reductions in consonant and vowel performance with decreasing speech level.

Figure 7 replots the mean sound field data for SPEAK and CIS listeners from Fig. 6, together with the mean sound field medium loud levels (MLLs) (left and center panels). MLLs for the speech frequencies (500 to 4kHz) were similar for the two groups, averaging 66.4 dB SPL for SPEAK listeners and 61.1 dB SPL for CIS listeners. Six of seven CIS participants did not reach medium loudness at 250 Hz at the output limits of the audiometer (83 dB SPL) and MLLs were arbitrarily assigned to this value. Thus, the data point for CIS participants at 250 Hz underestimates actual MLLs.

Dynamic ranges computed by subtracting sound field thresholds from sound field MLLs for speech-frequency stimuli (500 Hz to 4kHz) averaged 23.9 dB SPL (range, 12.1 to 32.9 dB) for SPEAK participants and 33.6 dB SPL (range, 30 to 39.3 dB) for CIS participants. In most cases, these dynamic ranges were considerably smaller than the nominal acoustic dynamic range implemented in participants' clinical maps, namely, 26 to 30 dB for SPEAK participants and 50 to 60 dB for CIS participants. Mean dynamic range data for SPEAK and CIS participants are plotted in the right panel of Fig. 7.

Figure 8 shows individual listeners' normalized HINT sentence scores at 50 dBA plotted as functions of sound field threshold (left panel) and sound field dynamic range (right panel). Sentence scores for the 50-dBA condition are normalized by expressing them as a percentage of each listeners' score at 70 dBA, so that relative performance at 50 dBA can be compared across listeners. Sound field measures are the mean values obtained across speech frequencies (500 to 4kHz) in each individual. It can be seen that normalized sentence scores showed considerable variability across listeners at the 50-dBA level, with CIS listeners generally maintaining higher scores than SPEAK listeners. The left scatterplot in Fig. 8 indicates that more than half of this variability may be accounted for by sound field thresholds (R^2 = 0.566, $p < 0.005$). This finding supports the notion that audibility is the primary factor underlying reduced speech recognition at moderate and low presentation levels. The right scatterplot in Fig. 8 indicates that normalized sentence scores at 50 dBA are also positively related to sound field dynamic range ($R^2 = 0.405$, $p < 0.005$). However, this association may simply reflect an inverse correlation

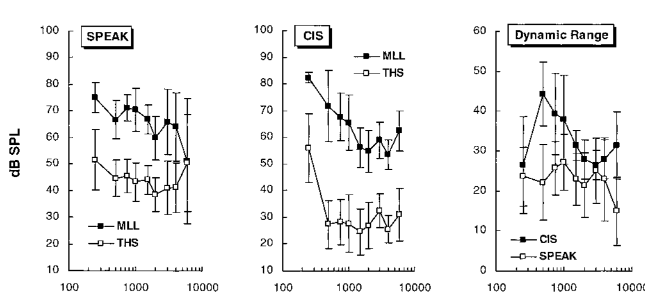

Figure 7. Mean sound field thresholds and medium loud levels (MLLs) for SPEAK and CIS listeners (left and middle panels, respectively), and mean dynamic ranges for the same listeners (right panel). Error bars indicate ± 1 SD.

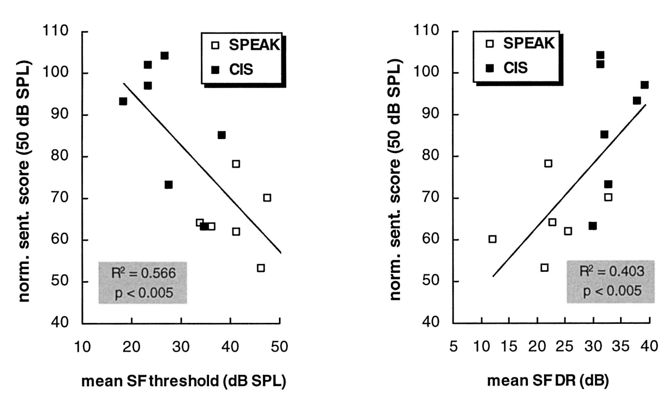

Figure 8. Normalized sentence scores obtained at 50 dBA plotted as functions of sound field threshold (left panel) and sound field dynamic range (right panel). Individual sound field threshold and dynamic range values represent mean values for frequencies between 500 and 4000 Hz.

between sound field threshold and sound field dynamic range ($R^2 = 0.554$, $p < 0.005$).

Effect of Sensitivity Setting

One way in which SPEAK listeners could potentially compensate for the limited acoustic dynamic range of the Nucleus 22 device is to adjust the sensitivity control on their processor as speech changes in level. As indicated earlier, the sensitivity control determines input gain: At high sensitivity settings, gain is increased with the result that softer sounds are mapped higher into the electric dynamic range. This improves audibility for low-level sounds. However, it also reduces the upper limit of the acoustic dynamic range, potentially compressing amplitude peaks in speech that may contribute to speech recognition. Further, because a greater proportion of speech peaks are presented at higher current levels within the electric dynamic range, speech at moderate acoustic amplitudes may become uncomfortably loud to the listener.

To determine whether our SPEAK participants might have achieved better overall speech recognition performance by using a different sensitivity setting on their device, we measured consonant and vowel recognition in two participants (N28 and N30) for three different sensitivity settings. These settings (2, 4 and 6) sampled most of the range available (0 to 8). Procedures were the same as those described earlier, except that only two blocks of data were obtained at each presentation level. Results, expressed as percent-correct phonemes, are shown in Figure 9.

The data in Figure 9 show that performance for soft-to-moderate level stimuli (40 to 60 dBA) improved substantially as the sensitivity control on the Spectra speech processor was increased from a setting of 2 to a setting of 6. Stimuli remained inaudible or barely audible at the lowest presentation level (30 dBA) when sensitivity was set to 2 or 4. In contrast, stimuli sometimes became too loud to permit testing at high stimulus levels (60 to 70 dBA) when sensitivity was set to 6. This accounts for the missing data points at 60 and 70 dBA in the figure. Thus, although increasing the sensitivity setting improved recognition of phonemes presented at very soft to moderately soft levels (30 to 50 dBA), it resulted in excessive loudness for moderate and slightly loud speech (60 to 70 dBA). This suggests that adjusting the sensitivity setting may allow SPEAK listeners to optimize their speech recognition performance when there is a single speech source of relatively constant loudness. However, it would probably not be effective if the listener were attempting to follow several voices at different levels. It is also impractical for implant users (especially children) to continuously adjust the sensitivity dial on their processors as the listening environment changes.

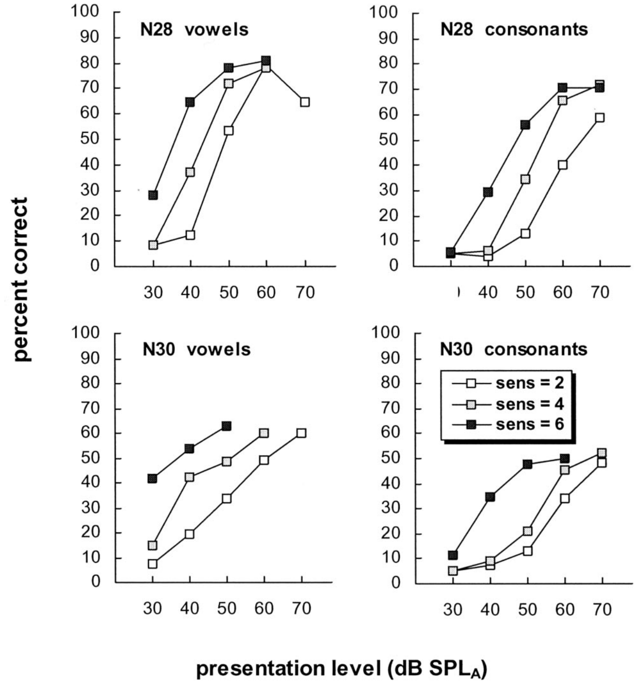

Figure 9. Performance-intensity functions for vowel and consonant recognition in two SPEAK listeners, for three different speech-processor sensitivity settings.

DISCUSSION

The primary finding of this study is that speech recognition in quiet is strongly level dependent for cochlear implant listeners, with performance declining sharply as presentation level is reduced from conversational or slightly loud levels (60 to 70 dBA) to softer levels (30 to 40 dBA). This contrasts with the situation in normal acoustic hearing in which high levels of speech recognition are maintained over a wide range of presentation levels. The present data suggest that the decrements shown by cochlear implant listeners at low levels are primarily the result of limited audibility. Both Nucleus SPEAK and Clarion CIS listeners in this study had sound field thresholds more than 20 dB poorer than those expected in normal-hearing listeners, and average speech-frequency thresholds for SPEAK listeners were 15 dB poorer than those for CIS listeners. Decrements in speech recognition at low presentation levels reflected these differences in audibility: P-I functions for SPEAK participants clearly rolled off more rapidly with decreasing speech level than those for CIS participants, and CIS participants showed stronger level effects than normal-hearing listeners.

Presentation level effects for Nucleus 22 participants in this study were similar to those observed by Skinner, Holden, Holden, Demorest, and Fourakis (1997) over the range of levels which they evaluated (50 to 70 dBA). Specifically, both studies showed strong level effects for consonants, words and sentences but relatively small effects for vowels. These results reflect the fact that consonant cues are lower in intensity than vowel cues and, thus, more susceptible to audibility limitations. Sentence recognition in both studies showed sharp level effects, suggesting that consonant perception is critical to listeners' understanding of higher-context speech materials.

Differences in average sound field thresholds exhibited by SPEAK and CIS listeners can be attributed at least partly to differences in the size of the acoustic dynamic range encoded by the Nucleus 22 and Clarion v1.2 devices. In this study, the acoustic dynamic range used by participants' speech processors varied from 26 to 30 dB for SPEAK participants and 50 to 60 dB for CIS participants. When listeners adjusted the sensitivity or volume control on their speech processor to yield comfortable loudness for 60-dBA speech, speech presented at 30 dBA was no longer fully audible. This was especially true among SPEAK participants, who were sometimes unable to hear any of the stimuli presented at the 30 dBA level. However, consideration of the individual data for CIS listeners indicates that IDR could not have been the only factor governing sound field thresholds and speech recognition at low presentation levels: Subjects C08 and C09 used 60 dB IDRs, yet had sound field thresholds higher than 40 dB SPL and showed greater reductions in speech scores at 30 and 40 dB SPL than the other CIS participants.

Other factors that may have affected individual listeners' results are the minimum and maximum stimulation settings used in their maps, and the sensitivity settings used during testing. As indicated in the Introduction, map minimum levels should ideally be set high enough to eliminate regions of very slow loudness growth near threshold. This insures that soft sounds are mapped into an audible portion of the electric dynamic range, thereby allowing them to contribute to speech recognition. Similarly, map maximum levels should be adjusted so that loud sounds are never too loud or distorted. Although the mapping procedures used in this study attempted to set map minimum and maximum levels according to these guidelines, the sensitivity settings selected by some individual SPEAK subjects suggest that minimum and/or maximum stimulation levels may not have been optimal. For example, three SPEAK participants (N28, N30 and N32) chose very low sensitivity settings of 1.5 and 2 to make speech at 60 dBA comfortably loud. It is possible that the maximum stimulation levels specified in the listeners' maps were too high, and that listeners selected low sensitivity control settings so that 60 dBA speech was not too loud. This possibility is consistent with the results shown in Fig. 9. That is, when sensitivity was set to 4, listeners noted that 70-dBA speech was too loud, even though this sensitivity setting allowed them to achieve better speech recognition at 40 dBA than a sensitivity setting of 2. Two other SPEAK participants, N12 and N14, chose unusually high sensitivity settings (5.5 and 5, respectively), suggesting that their map minimum levels were set too low and that higher sensitivity settings were needed to make speech loud enough at 60 dBA. Although their sound field thresholds were among the lowest of the SPEAK participants, they had some of the poorest sentence scores at 30, 40 and 50 dBA (Fig. 5). It is likely that minimum stimulation levels were also set too low for CIS participants C08 and C09 and that these listeners would have achieved lower sound field thresholds and higher speech recognition scores at 30 and 40 dBA if minimum stimulation levels had been raised. In general, then, optimization of minimum and maximum stimulation levels in the present study may have resulted in lower sound field thresholds and better speech recognition at low presentation levels, especially among SPEAK participants. The importance of these factors has been addressed by Skinner et al. (1999) and Holden, Skinner, Holden and Demorest (2002). Figure 1 in Holden et al. (2002) demonstrates that optimum settings for minimum and maximum stimulation levels may differ significantly from the thresholds and maximum acceptable loudness levels measured on individual electrodes.

It should be noted that the Nucleus-24 SPRINT speech processor appears to support lower sound field thresholds than the Nucleus 22 Spectra processor, even though it uses a similar, 30-dB IDR. After optimizing speech processor settings, eight Nucleus-24 users tested in a study by Skinner, Holden, Whitford, Plant, Psarros, and Holden (2002) exhibited a mean sound field threshold of 27 dB SPL (500 to 4000 Hz; Skinner, personal communication). This value is similar to that measured for Clarion CIS listeners in the present study and is approximately 15 dB better than that measured for our Nucleus 22 SPEAK listeners. Although Skinner et al. did not measure speech recognition at low presentation levels in these subjects, it is likely that their performance would be significantly better than that exhibited by SPEAK subjects in the present study. Improved audibility with the SPRINT processor may be due to its use of fully digital circuits and to the addition of a volume control that is not present on the Spectra processor. Implementation of adaptive dynamic range optimization with the SPRINT processor may also have a substantial impact on Nucleus-24 listeners' ability to recognize soft speech (James, Blamey, Martin, Swanson, Just, & Macfarlane, 2002).

CONCLUSIONS

Both the Nucleus 22 SPEAK and Clarion v1.2 CIS speech processing strategies support high levels of speech recognition when speech is presented at normal or raised conversational levels. However, performance drops substantially with reductions in presentation level, especially for SPEAK users. This represents a significant limitation for device benefit and may be especially problematic for young children who are in the process of acquiring language. Speech processors that use wide acoustic dynamic ranges may provide better audibility for speech sounds and may thereby improve speech recognition at low presentation levels in quiet. Audibility and recognition of low-level speech may also be affected by a number of other factors, including minimum and maximum electrical stimulation levels, the sensitivity control setting and the characteristics of speech-processor circuitry.

ACKNOWLEDGMENTS:

This research was supported by NIDCD Grant No. DC00110 and the Lions 5M International Hearing Foundation. Preliminary findings were presented at the 1999 American Academy of Audiology annual convention (Miami, FL). Drs. James Hillenbrand and Dianne Van Tasell provided the digitized vowel and consonant stimuli, respectively, used in the study; John van Essen wrote the computer programs used to obtain phoneme confusion data; and Tanya Grann assisted with data collection. Peggy Nelson, Ben Munson, Arlene Carney, Margaret Skinner and two anonymous reviewers made many constructive comments on an earlier version of the manuscript. The authors thank Dr. Skinner for her insightful observations regarding the interpretation of results, many of which were incorporated into the Discussion. The authors extend special thanks to the 14 cochlear implant listeners who participated in the study.

Address for correspondence: Gail S. Donaldson, Ph.D., University of Minnesota, Department of Otolaryngology, MMC 396, 420 Delaware St., S.E., Minneapolis, MN 55455. E-mail: donal005@umn.edu.

Received January 11, 2002; accepted March 18, 2003

REFERENCES

Blamey, P. J., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., Larky, J., Lindstrom, B., Nedzelski, J., Peterson, A., Shipp, A., Staller, S., & Whitford, L. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology and Neuro-Otology*, *1*, 293–306.

Carhart, R., & Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders*, *24*, 330–345.

French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, *19*, 90–199.

Fu, Q-J., & Shannon, R. V. (1998). Effects of amplitude nonlinearity on speech recognition by cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *104*, 2571–2577.

Fu, Q-J., & Shannon, R. V. (1999). Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users. *Journal of the Acoustical Society of America*, *106*, L65–L70.

Fu, Q-J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electrical hearing. *Journal of the Acoustical Society of America*, *104*, 3586–3596.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.

Hochberg, I., Boothroyd, A., Weiss, M., & Hellman, S. (1992). Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear and Hearing*, *13*, 263–271.

Holden, L. K., Skinner, M. W., Holden, T. A., & Demorest, M. E. (2002). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear and Hearing*, *23*, 463–476.

James, C. J., Blamey, P. J., Martin, L., Swanson, B., Just, Y., & Macfarlane, D. (2002). Adaptive dynamic range optimization for cochlear implants: A preliminary study. *Ear and Hearing*, *23*, 29S–58S.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, *27*, 338–352.

Muller-Deile, J., Schmidt, B. J., & Rudert, H. (1995). Effects of noise on speech discrimination in cochlear implant patients. *Annals of Otology, Rhinology and Laryngology*, *104*(Suppl.166), 303–306.

Nilson, J., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, *95*, 1085–1099.

Seligman, P., & Whitford, L. (1995). Adjustment of appropriate signal levels in the Spectra 22 and mini speech processors. *Annals of Otology, Rhinology and Laryngology*, *104*(Suppl. 166), 172–175.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P. M., Staller, S. J., Shipp, D. B., Shallop, J. K., Everingham, C., Menapace, C. M., Arndt, P. L., Antogenelli, T., Brimacombe, J. A., Pijl, S., Daniels, P., George, C. R., McDermott, H. J., & Beiter, A. L. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. *American Journal of Otology*, *15*, 15–27.

Skinner, M. W., Holden, L. K., & Holden, T. A. (1997). Parameter selection to optimize speech recognition with the Nucleus implant. *Otolaryngology. Head and Neck Surgery*, *117*, 188–195.

Skinner, M. W., Holden, L. K., Holden, T. A., & Demorest, M. E. (1995). Comparison of procedures for obtaining thresholds and maximum acceptable loudness levels with the Nucleus cochlear implant system. *Journal of Speech and Hearing Research*, *38*, 677–689.

Skinner, M. W., Holden, L. K., Holden, T. A., & Demorest, M. E. (1999). Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. *Journal of Speech, Language and Hearing Research*, *42*, 814–828.

Skinner, M. W., Holden, L. K., Holden, T. A., Demorest, M. E., & Fourakis, M. S. (1997). Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants. *Journal of the Acoustical Society of America*, *101*, 3766–3782.

Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C., & Holden, T. A. (2002). Speech recognition with the Nucleus SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear and Hearing*, *23*, 207–223.

Sun, J. C., Skinner, M. W., Liu, S. Y., Wang, F. N. M., Huang, T. S., & Lin, T. (1998). Optimization of speech processor fitting strategies for chinese-speaking cochlear implantees. *Laryngoscope*, *108*, 560–568.

Van Tasell, D. J., Greenfield, D. G., Logemann, J. J., & Nelson, D. A. (1992). Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants. *Journal of the Acoustical Society of America*, *92*, 1247–1257.

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

Zeng, F-G., & Galvin, J. J. (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners. *Ear and Hearing*, *20*, 60–74.

Zeng, F-G., Grant, G., Niparko, J., Galvin, J., Shannon, R., Opie, J., & Segel, S. (2002). Speech dynamic range and its effect on cochlear implant performance. *Journal of the Acoustical Society of America*, *111*, 377–386.

Reference Notes

1 Wilson, B. (2000). *Sixth Quarterly Progress Report, Speech Processors for Auditory Prostheses*. Research Triangle Park, NC: Research Triangle Institute.

2 Allen, S. A., & Donaldson, G. S. (1999). Input dynamic range and speech perception in cochlear implant listeners. (Unpublished data).