

Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing

QIAN-JIE FU AND GERALDINE NOGAKI

Department of Auditory Implants and Perception, House Ear Institute, 2100 West Third Street, Los Angeles, CA 90057, USA

Received: 22 March 2004; Accepted: 15 September 2004; Online publication: 2 February 2005

ABSTRACT

The latest-generation cochlear implant devices provide many deaf patients with good speech recognition in quiet listening conditions. However, speech recognition deteriorates rapidly as the level of background noise increases. Previous studies have shown that, for cochlear implant users, the absence of fine spectro-temporal cues may contribute to poorer performance in noise, especially when the noise is dynamic (e.g., competing speaker or modulated noise). Here we report on sentence recognition by cochlear implant users and by normal-hearing subjects listening to an acoustic simulation of a cochlear implant, in the presence of steady or square-wave modulated speech-shaped noise. Implant users were tested using their everyday, clinically assigned speech processors. In the acoustic simulation, normal-hearing listeners were tested for different degrees of spectral resolution (16, eight, or four channels) and spectral smearing (carrier filter slopes of -24 or -6 dB/octave). For modulated noise, normal-hearing listeners experienced significant release from masking when the original, unprocessed speech was presented (which preserved the spectro-temporal fine structure), while cochlear implant users experienced no release from masking. As the spectral resolution was reduced, normal-hearing listeners' release from masking gradually diminished. Release from masking was further reduced as the degree of spectral smearing increased. Interestingly, the mean speech recognition thresholds of implant users were very close to those of normal-hearing

subjects listening to four-channel spectrally smeared noise-band speech. Also, the best cochlear implant listeners performed like normal-hearing subjects listening to eight- to 16-channel spectrally smeared noise-band speech. These findings suggest that implant users' susceptibility to noise may be caused by the reduced spectral resolution and the high degree of spectral smearing associated with channel interaction. Efforts to improve the effective number of spectral channels as well as reduce channel interactions may improve implant performance in noise, especially for temporally modulated noise.

Keywords: noise susceptibility, cochlear implants, spectral resolution, spectral smearing, gated noise

INTRODUCTION

With the introduction of the cochlear implant (CI) device, hearing sensation has been restored to many profoundly deaf patients. Many postlingually deafened patients, fitted with the latest multichannel speech processors, perform very well in quiet listening situations. However, speech performance deteriorates rapidly with increased levels of background noise, even for the best CI users. Understanding CI users' susceptibility to noise remains a major challenge for researchers (e.g., Dowell et al. 1987; Hochberg et al. 1992; Kiefer et al. 1996; Müller-Deiler et al. 1995; Skinner et al. 1994), and is an important step toward improving CI users' performance in noisy listening conditions.

Recent innovations in electrode design and speech processing have helped to improve CI users' performance in noise (Skinner et al. 1994; Kiefer et al. 1996). However, even with the latest technology, the

Correspondence to: Qian-Jie Fu • Department of Auditory Implants and Perception • House Ear Institute • 2100 West Third Street, Los Angeles, CA 90057, USA. Telephone: (213) 273-8036; fax: (213) 413-0950; email: qfu@hei.org

speech recognition of CI listeners is more susceptible to background noise than that of normal-hearing (NH) listeners, especially when the interfering sound is competing speech or temporally modulated noise (Müller-Deiler et al. 1995; Nelson et al. 2003). Nelson et al. (2003) measured sentence recognition in the presence of both steady and modulated speech-shaped noise in both CI and NH listeners. Not surprisingly, NH listeners obtained significant release from masking from modulated maskers, especially at the 8-Hz masker modulation frequency, consistent with previous findings that NH listeners may take advantage of the temporal gaps in the fluctuating maskers (Bacon et al. 1998; Gustafsson and Arlinger 1994). In contrast, CI users experienced very little release from masking from modulated maskers; speech recognition was more adversely affected at syllabic modulation rates (2–4 Hz). The authors argued that CI users' speech processing strategies and/or lack of spectro-temporal details in the processed speech did not allow for the release from masking with modulated maskers observed with NH listeners. This observation also applies to subjects with sensorineural hearing loss when listening to amplified speech (e.g., Bacon et al. 1998; Festen and Plomp 1990; Takahashi and Bacon 1992). Audibility alone cannot explain the additional masking experienced by listeners with sensorineural hearing loss (Eisenberg et al. 1995).

These previous studies suggest that diminished spectral resolution is responsible for increased susceptibility to noise (Fu et al. 1998; Nelson et al. 2003; Nelson and Jin 2004; Qin and Oxenham 2003). For hearing-impaired (HI) subjects listening to amplified speech, broadened auditory tuning curves can reduce the spectral resolution (e.g., Horst 1987). Several factors limit CI users' available spectral resolution: (1) the number of implanted electrodes (presently, 16–22, depending on the implant device); (2) the uniformity and health of surviving neurons and their proximity to the implanted electrodes; and (3) the amount of current spreading from the stimulating electrodes. While the first factor describes the limited spectral resolution transmitted by the implant device, the second and third factors describe the amount of spectral detail received by CI users (which may differ greatly between patients). When these spectral details are smeared (because of channel/electrode interactions), CI users' effective spectral resolution can be further reduced, relative to the spectral information transmitted by the implant device. The effects of spectral smearing on speech recognition in noise have been well documented (e.g., ter Keurs et al. 1992, 1993; Baer and Moore 1993, 1994; Nejime and Moore 1997; Boothroyd et al. 1996). In general, reduced spectral resolution be-

cause of spectral smearing will cause a significant performance drop in quiet and worsen further in noise.

The present study investigates whether CI users' increased noise susceptibility, especially to dynamically changing noise, is a result of the reduced spectral resolution caused by the limited number of spectral channels and/or channel interactions across the simulated electrodes. Speech recognition thresholds (SRTs) in the presence of steady-state or square-wave modulated speech-shaped noise were measured in CI users as well as in NH subjects listening to an acoustic CI simulation. Cochlear implant users were tested using their everyday, clinically assigned speech processors. Normal-hearing listeners were tested for different degrees of spectral resolution (16, eight, or four channels) and spectral smearing (–24 or –6 dB/octave carrier band filters). The data from NH subjects for these spectral resolution and smearing conditions were compared to those from CI users.

METHODS

Subjects

Ten CI users and six NH subjects participated in the experiment. Cochlear implant subjects were 35–70 years old. Nine were postlingually deafened. One (M1) was deafened at seven years of age. All were native speakers of American English and all had at least three months experience with their CI device and mapping strategy. Four subjects (S1–S4) used the Nucleus-22 device, one subject used the Clarion I device, and five subjects used the Clarion CII device. All were considered “good users” of the implant device, in that they were able to achieve near-perfect HINT sentence recognition in quiet with their everyday speech processors. Poor users were not selected for this study to avoid floor effects. All implant participants had extensive experience in speech recognition experiments. Table 1 contains relevant information for the 10 CI subjects. The six NH listeners (three males and three females) were 26–37 years old. All NH subjects had pure tone thresholds better than 15 dB HL at audiometric test frequencies from 250 to 8000 Hz and all were native speakers of American English.

Cochlear implant speech processor settings

All CI listeners were tested using their everyday, clinically assigned speech processors. Information regarding each CI subject's implant device and speech processor parameters is shown in Table 1. Sentence recognition was conducted in free field in a double-walled sound treated booth (IAC). The signal

TABLE 1

Subject information for 10 cochlear implant listeners who participated in the present study							
Subject	CI	Speech processing strategy	Input frequency range (kHz)	Number of channels	Stimulation rate per electrode (ppse)	Age	Implant use (years)
S1	Nuc22	SPEAK	0.15–10.15	22	250	45	10
S2	Nuc22	SPEAK	0.15–10.15	22	250	61	12
S3	Nuc22	SPEAK	0.15–10.15	22	250	51	7
S4	Nuc22	SPEAK	0.15–10.15	22	250	60	10
M1	Med-El	CIS+	0.2–8.5	12	600	39	3
C1	CII/HiFocus	HiRes	0.33–6.66	16	4640	50	4
C2	Clarion I	CIS	0.25–6.8	8	812.5	58	5
C3	CII	HiRes	0.33–6.66	16	4640	70	4
C4	CII	CIS	0.25–6.8	8	812.5	35	0.25
C5	CII	CIS	0.25–6.8	8	812.5	54	2

was presented at 65 dBA *via* a single loudspeaker (Tannoy Reveal); subjects were seated 1 m from the loudspeaker and instructed to face the loudspeaker directly. Subjects were instructed to use their everyday volume and microphone sensitivity settings and disable any noise reduction settings in their speech processors. Once set, subjects were asked to refrain from manipulating their processor settings during the course of the experiment.

Signal processing for acoustic simulation

For NH listeners, speech and noise were processed by an acoustic simulation of a CI. Noise-band vocoder speech processing was implemented as follows. The input acoustic signal was band-pass filtered into several broad frequency bands (16, eight, or four,

depending on the test condition) using fourth-order Butterworth filters (-24 dB/octave). Table 2 shows the cutoff frequencies of the analysis bands for the three spectral resolution conditions. The amplitude envelope was extracted from each band by half-wave rectification and low-pass filtering (fourth-order Butterworth) with a 160-Hz cutoff frequency. The extracted envelope from each band was used to modulate a wideband noise, which was then spectrally limited by a bandpass filter corresponding to the analysis band filter. Depending on the spectral smearing test condition, the slope of the carrier band filters was either -24 dB/octave (which matched the analysis band filter slope and simulated a CI with little channel interaction) or -6 dB/octave (which simulated a CI with significantly more channel interaction). The outputs from each band were

TABLE 2

Analysis and carrier band cutoff frequencies used in noise-vocoder cochlear implant simulation								
Cutoff frequency (Hz)								
Channel	Lower	Upper	Channel	Lower	Upper	Channel	Lower	Upper
1	200	272	1	200	359	1	200	591
2	272	359						
3	359	464	2	359	591			
4	464	591						
5	591	745	3	591	931	2	591	1426
6	745	931						
7	931	1155	4	931	1426			
8	1155	1426						
9	1426	1753	5	1426	2149	3	1426	3205
10	1753	2149						
11	2149	2627	6	2149	3205			
12	2627	3205						
13	3205	3904	7	3205	4748	4	3205	7000
14	3904	4748						
15	4748	5768	8	4748	7000			
16	5768	7000						

summed and presented to NH listeners in free field *via* a single loudspeaker (Tannoy Reveal) at 65 dBA in a double-walled, sound-treated booth (IAC).

Figure 1 shows an example of the acoustic processing used to simulate spectral smearing in a CI. The left panel shows the averaged short-term spectra for the vowel /i/, processed by a 16-channel noise-band vocoder. The steep carrier filter slope (-24 dB/octave; dotted line) clearly preserves the formant structure, including formant frequencies and the valley between formants. While the formant frequencies are still well preserved by the shallow carrier filter slope (-6 dB/octave; solid line), the level difference between formant peaks and spectral valleys is significantly reduced. The right panel shows the averaged short-term spectra for the vowel /i/, processed by a four-channel noise-band vocoder. With four spectral channels and steep filter slopes, the formant structure is less clear than with the 16-channel processor with steep filter slopes; the second formant is much less distinct. With four spectral channels and shallow filter slopes, the level difference between formant peaks and spectral valleys is further reduced. In general, the number of spectral channels will mainly affect the peak formant frequencies, while the slope of the carrier filters will mainly affect the level difference between formant peaks and spectral valleys. The smearing method used in the present study aims at simulating electrode interactions in CIs and is slightly different from smearing methods used in previous studies (e.g., Baer and Moore 1994; Boothroyd et al. 1996). Nevertheless, the change of acoustic characteristics (e.g., formant

patterns) in the frequency domain is comparable for all these smearing approaches. Note, however, that these various approaches to spectral smearing may not be equivalent in the time domain.

Test materials and procedure

Speech materials consisted of sentences taken from the Hearing in Noise Test (HINT) sentence set (Nilsson et al. 1994). This set consists of 260 sentences, spoken by one male talker. Speech-shaped noise (either steady-state or gated at various interruption rates) was used as a masker. Gated noise frequencies ranged from 1 to 32 Hz, resulting in noise bursts that ranged in duration from 16 ms (32 Hz) to 500 ms (1 Hz). Gating was implemented using 50% duty cycles and 100% modulation depths; all gates had 2-ms cosine squared ramps. The onset of noise was 500 ms before the sentence presentation and the offset of noise was 500 ms after the end of the sentence presentation; masker rise and fall times were 5 ms.

During testing, NH subjects listened to combined speech and noise processed by CI simulations (or unprocessed speech and noise for baseline measures); CI subjects listened to combined speech and noise (unprocessed) through their everyday, clinically assigned speech processors. Speech and noise were mixed at the target signal-to-noise ratio (SNR) prior to processing. The SNR was defined as the difference, in decibels, between the long-term root mean square (RMS) levels of the sentence and the noise. Because the RMS level of gated noise was calculated for the whole cycle, the peak RMS level of gated noise was 3 dB higher than that of steady noise at the same SNR. After mixing the sentence and the noise at the target SNR, the overall presentation level was adjusted to 65 dBA to maintain comfortable listening levels at all SNRs. For NH listeners, the combined speech and noise was then processed by the CI simulation and output to a single loudspeaker. For CI listeners (and NH subjects listening to unprocessed speech), the unprocessed speech and noise was output to a single loudspeaker.

Subjects were tested by using an adaptive SRT procedure. The SRT has been defined as the SNR that produces 50% correct sentence recognition (Plomp 1986). In the Plomp study, sentence recognition was measured at various fixed SNRs, producing a psychometric function that was sigmoidal in shape. Performance was relatively unaffected until the SNR was sufficiently low, beyond which recognition dropped steeply. The SNR corresponding to 50% correct whole-sentence recognition along the psychometric function was defined as the SRT. Because of the limited sentence materials and numerous test con-

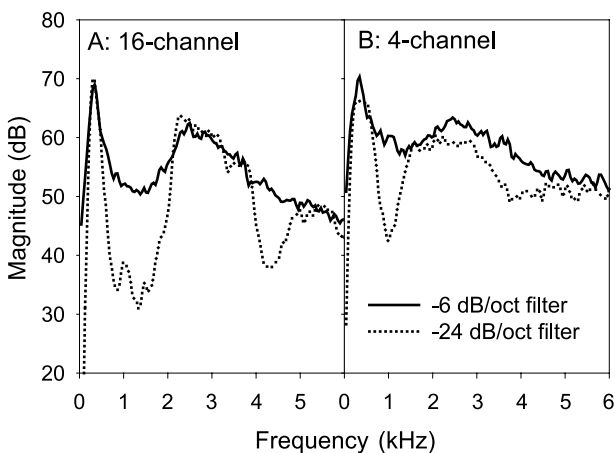


FIG. 1. Averaged short-term spectra for noise-band vocoder speech processing used in the acoustic cochlear implant simulation for the vowel /i/. The left panel shows processing by 16 channels and the right panel shows processing by four channels. The dotted lines show the short-term spectra without smearing (-24 dB/octave filter slopes) and the solid lines show the short-term spectra with smearing (-6 dB/octave filter slopes).

ditions, a modified method was used to measure the SRT in the present study. The SNR was adapted according to the correctness of response. Because the HINT sentences were easy (all subjects were capable of nearly 100% correct in quiet), sentence recognition was very sensitive to the SNR; near the SRT, most subjects were able to either immediately recognize the sentence, or not. Thus SNR adjustments were made along the steep slope of Plomp's psychometric function. In the present study, the SRT was measured by using an adaptive 1-up/1-down procedure (Nilsson et al. 1994). When the subject repeated every word of the sentence correctly, the SNR was reduced by 2 dB; if the subject did not repeat every word correctly, the SNR was increased by 2 dB. Within each test run, the mean of the final eight reversals in the SNR was defined as the SRT (the SNR needed to produce 50% correct sentence recognition). For all data shown, a minimum of two runs were averaged for each condition. A potential problem with the adaptive SRT procedure is intrasubject variability between test runs, which can make the threshold measurement somewhat noisy. To reduce this source of variability, additional runs were conducted when the difference between two runs was more than 6 dB. For most CI subjects, four or more runs per condition were collected.

For NH listeners, the -24 dB/octave carrier filter slope conditions were tested before the -6 dB/octave conditions at all spectral resolution and noise conditions. The spectral resolution and noise condition were otherwise randomized and counterbalanced across subjects. For CI listeners, the noise conditions were randomized and counterbalanced across subjects.

RESULTS

Figure 2 shows mean SRTs as a function of noise conditions for NH and CI subjects listening to unprocessed speech in noise. The filled circles represent the data from the NH listeners while the open circles represent the data from the CI listeners. For NH listeners, the poorest performance was observed with steady-state noise (SRT = -4.6 dB). Performance improved with gated noise, and the peak performance was observed with 4-Hz gated noise (SRT = -22.2 dB). Performance fell to -8.5 dB as the modulation frequency was increased from 4 to 32 Hz. A one-way repeated measures analysis of variance (ANOVA) revealed that NH listeners' SRTs with unprocessed speech were significantly affected by the noise conditions [$F(6, 35) = 24.255, p < 0.001$]. A *post hoc* Tukey's test showed that better speech

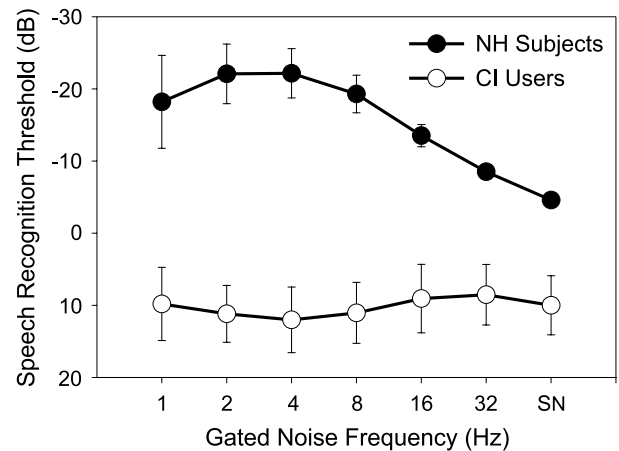


FIG. 2. Mean speech recognition thresholds as a function of noise conditions in normal-hearing listeners and cochlear implant patients for the original unprocessed speech. The filled circles represent the data from the NH listeners while the open circles represent the data from the CI listeners. The error bars represent one standard deviation.

recognition was observed for gated frequencies of 16 Hz or less, relative to steady-state noise.

The mean SRT for CI patients was $+10.0$ dB in steady-state noise, about 14.6 dB worse than that of NH listeners. In contrast to NH listeners' results, the poorest performance was observed with 4-Hz modulated noise (SRT = 12.0 dB), while peak performance was observed for 32-Hz modulated noise (SRT = 8.5 dB). The largest difference between the SRTs for various noise conditions was only 3.5 dB for CI users, as compared to 17.6 dB for NH listeners. An ANOVA test revealed no significant main effect of noise conditions on the SRTs [$F(6, 63) = 0.789, p = 0.582$], most likely as a result of the high variability across CI patients (individual subject data are provided in Fig. 3). Within-subject paired Student's *t*-tests did reveal that CI users' SRTs were significantly worse for the 2-Hz [$t(9) = 2.75, p = 0.023$] and 4-Hz gated noise conditions [$t(9) = 3.35, p < 0.01$], and significantly better for 32 Hz [$t(9) = 3.45, p < 0.01$], relative to steady-state noise. Defining masking release as the SRT difference between gated noise and steady-state noise, ANOVA tests showed a small but significant effect of modulation frequency on the amount of masking release in CI patients [$F(5, 54) = 4.765, p = 0.001$].

CI users' SRTs were highly variable, even among users of the same implant device and speech processing strategy. Figure 3 shows the mean performance of each CI device (solid lines), with individual CI users' scores shown as unfilled symbols. The best performer in each group is represented by the open circles. The best performer among all CI subjects used the Clarion CII device with HiRes (subject C1); the poorest performer used the Clarion CII device with

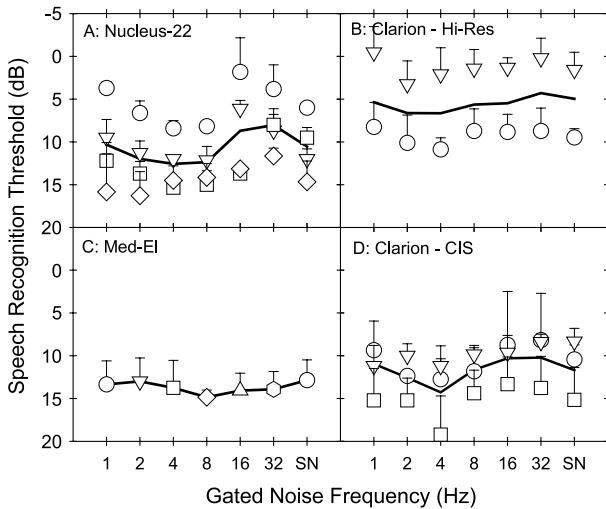


FIG. 3. Individual and mean SRTs for CI subjects, grouped by device and speech processing strategy. **A.** Four patients with Nucleus-22 device; **B.** two patients with Clarion CII/Hi-Res strategy; **C.** one patient with Med-EI device; **D.** three patients with Clarion/CIS strategy. The error bars represent one standard deviation.

CIS (subject C5). It should be noted there was only one MedEI user in the CI group (subject M1). Unfortunately, there were not enough users per device or processing strategy type to allow formal statistical analysis of any implant device differences in performance.

Figure 4 shows the mean SRTs as a function of noise conditions in NH subjects listening to four-, eight-, and 16-channel noise-band speech. Panel A shows the data with -24 dB/octave carrier filter slope while panel B shows results with -6 dB/octave carrier filter slope. The mean SRTs for NH and CI subjects with unprocessed speech are shown by the dotted and dashed lines, respectively. For the steep carrier filter slope conditions (-24 dB/octave), reduced spectral resolution negatively affected NH listeners' SRTs. Relative to unprocessed speech, SRTs were significantly worse for the 16-channel processor with -24 dB/octave carrier filter slopes (which was the highest spectral resolution with the least channel interaction simulated in the study). The shift in SRTs was different for various noise conditions, ranging from 5.4 dB for steady-state noise to 13.7 dB for the 4-Hz modulation frequency (relative to unprocessed speech). When the spectral resolution was further reduced, SRTs worsened at all noise conditions. Similar to the 16-channel processor, for the eight- and four-channel processors, the minimum shift in SRTs occurred for the steady-state noise condition, while the peak shift occurred for the 2- or 4-Hz gated noise condition. A two-way ANOVA showed significant main effects of noise conditions [$F(6,140) = 34.633$, $p < 0.001$] and spectral channels [$F(3, 140) =$

265.153, $p < 0.001$], as well as significant interaction between noise conditions and spectral channels [$F(18, 140) = 4.576$, $p < 0.001$]. The NH subjects' SRTs for four-channel speech (-24 dB/octave carrier filter slopes) were significantly better than those of CI users, across all noise conditions [$F(1, 98) = 43.486$, $p < 0.001$].

When the spectrum was further smeared by the shallow -6 dB/octave carrier filter slope (Fig. 4B), the SRTs worsened for all spectral resolution conditions, relative to performance with the steep carrier filter slope (-24 dB/octave). For the -6 dB/octave filter condition, the SRT shift relative to the -24 dB/octave filter condition was 1.6 dB (16-channel), 3.3 dB (eight-channel), and 3.8 dB (four-channel) in steady-state noise. For gated noise, the maximum SRT shift relative to performance with the steep carrier filter slope was 11.3, 12.3, and 8.2 dB for 16-channel, eight-channel, and four-channel versions, respectively. A three-way ANOVA test (with spectral resolution, carrier filter slopes, and noise conditions as factors) showed significant main effects for spectral resolution [$F(2, 210) = 320.49$, $p < 0.001$], carrier filter slopes [$F(1, 210) = 297.89$, $p < 0.001$], and noise conditions [$F(6,210) = 14.87$, $p < 0.001$]. The ANOVA also showed significant interactions between the spectral resolution and noise conditions [$F(12, 210) = 3.00$, $p = 0.001$] and between carrier filter slopes and noise conditions [$F(6, 210) = 4.94$, $p < 0.001$], but no significant interaction between the spectral resolution and filter slopes [$F(2,210) = 1.48$, $p = 0.229$]. There was no significant interaction between all three factors [$F(12, 210) = 1.113$, $p = 0.351$].

Mean CI performance was comparable to that of NH subjects listening to four spectrally smeared channels (-6 dB/octave filters). The best CI user's performance was comparable to that of NH subjects

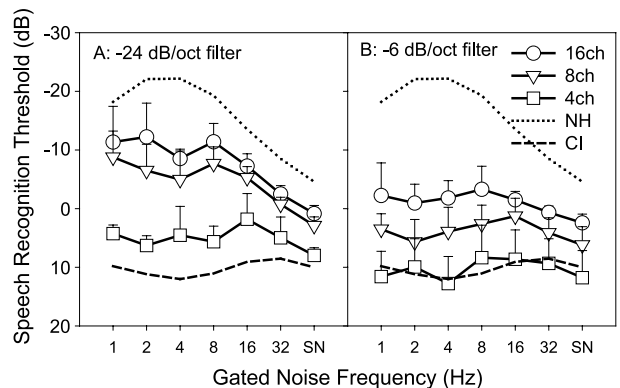


FIG. 4. Speech recognition thresholds as a function of gated noise frequency in NH listeners for 16-, eight-, and four-channel noise-band speech and unprocessed speech. **A.** Steep carrier filter slope (-24 dB/octave). **B.** Shallow carrier filter slope (-6 dB/octave). The error bars represent one standard deviation.

listening to 8–16 spectrally smeared channels, or four spectrally independent channels (−24 dB/octave). The poorest CI user's performance was sometimes worse than that of NH subjects listening to four spectrally smeared channels.

DISCUSSION

The data from the present study clearly demonstrate CI patients' high susceptibility to background noise. The results from NH subjects suggest that a combination of poor frequency resolution and channel interaction contribute strongly to CI users' difficulty in noisy listening situations, especially temporally modulated noise.

Regardless of the type of noise (e.g., steady-state or modulated), CI users' speech recognition in noise is negatively affected by the loss of spectro-temporal fine structure caused by the limited number of electrodes and spectral channels in the speech processing strategy, consistent with the previous results (Fu et al. 1998; Nelson et al. 2003; Nelson and Jin 2004; Qin and Oxenham 2003). Fewer numbers of channels degrade the spectral resolution, and speech recognition becomes difficult in noise. However, the number of spectral channels transmitted by the speech processor and implant device may not completely account for the poorer performance of CI users in noise, especially temporally modulated noise. Listening to four spectrally independent channels, NH subjects' mean SRTs were significantly better than that of CI users. Only when the spectral cues were smeared by the shallow filter slopes (−6 dB/octave) did NH listeners' performance become comparable to that of CI users. For all spectral resolution and noise conditions, NH listeners' performance worsened with the shallow carrier filter slopes. There was no release of masking with the modulated maskers for four- or eight-channel spectrally smeared speech. While mean CI performance was comparable to that of NH subjects listening to four spectrally smeared channels, the best CI user's performance was comparable to that of NH subjects listening to eight spectrally smeared channels, or four spectrally independent channels. This result suggests that, although more than four spectral channels are transmitted by the implant device (the CI devices tested transmitted between eight and 22 channels, depending on the manufacturer and the strategy), CI users may experience a high degree of channel interaction, thereby reducing their functional spectral resolution.

Consistent with previous studies' results, NH subjects experienced significant release from masking with the gated maskers when listening to the unpro-

cessed speech. The greatest amount of masking release (17.6 dB) was obtained with the 4-Hz masker, while the least amount (3.9 dB) was with the 32-Hz masker. With steady-state noise, −4.6 dB SNR was required to obtain 50% correct whole-sentence recognition. However, with 4-Hz gated noise, −22.2 dB SNR produced 50% correct sentence recognition. This suggests that NH subjects were able to “listen in the dips” of the masker envelope and better understand speech in noise. Even with reduced spectral resolution, NH subjects continued to experience significant amounts of masking release with the modulated maskers when the carrier filter slopes were steep. However, the amount of masking release was highly dependent on the degree of spectral resolution, i.e., masking release diminished as the number of spectral channels decreased. When spectral smearing was introduced to the CI simulation (−6 dB/octave carrier filter slopes), significant masking release was observed only for the 16-channel condition.

For CI subjects, the gated noise provided very little release from masking. The performance difference between steady-state and gated noise (at any interruption rate) was relatively small. The present findings were consistent with the results of Nelson et al. (2003) in that there was slight improvement in recognition at very high (32 Hz) interruption rates and a slight deficit in performance at low rates (4 Hz). It is possible that the 4-Hz gated masker, being temporally similar to syllabic speech rates, produced increased interference of the speech signal. Speech may even have been “fused” with the gated noise at the rate, in a kind of modulation detection interference across temporally similar envelopes. Because CI users lack the spectro-temporal fine structure to segregate the speech and noise sources, speech and noise may have been perceptually fused at this low modulation rate (e.g., Hall and Grose 1991; Yost 1992).

Although there were not enough CI users to allow for direct and fair comparison between implant devices and speech processing strategies, there seemed to be little difference between devices, at least in terms of CI users' release from masking. While there were individual performance differences in terms of absolute recognition scores, outside of the occasional masking release at various modulation frequencies, there were no clear differences in trends across implant devices. Thus the specific processing characteristics of the implant devices—including speech processing strategy, the number and location of implanted electrodes, the acoustic–electric frequency allocation, the acoustic input dynamic range and compression algorithms, the electric dynamic range, and amplitude mapping functions, etc. (all of which varied significantly across implant devices and individual implant users)—did not seem to account

for CI users' generally poor performance in modulated noise. Another important observation is that there was as much within-device variability as between-device variability. These results strongly suggest that the variability in masking release seen among implant users is individual-based rather than device-based.

Differences between the various implant devices in terms of acoustic input compression algorithms also make cross-device comparisons difficult, especially when using an adaptive SRT method. Differences between the "front-end" compressors of various devices may cause the effective SNR to be very different from the nominal SNR presented to CI subjects. Depending on the attack and release times, the compression threshold and ratio, the acoustic input dynamic range, and microphone sensitivity settings, CI users of different devices may have experienced very different SNRs at various points of the adaptive SRT test. The input SNR to a speech processor may be very different from the output of the front-end compressor. The subsequent acoustic-electric mapping may also compress the signal, further distorting the input SNR. Because it is quite difficult to quantify the effects of various compressors to the input signal for different devices, it is difficult to know whether these compressors may have significantly affected CI users' ability to obtain release from masking with modulated maskers.

SUMMARY AND CONCLUSIONS

The data from the present study demonstrate that CI users' sentence recognition is severely compromised in the presence of fluctuating noise even at fairly high SNRs (e.g., 6 dB SNR). As long as the spectral channels did not overlap, NH subjects were able to obtain significant release from masking, even when the spectral resolution was severely reduced. In contrast, when the carrier bands were smeared to simulate channel interaction (-6 dB/octave filter), NH listeners exhibited no significant release from masking, even when the spectral resolution was relatively high (eight channels). Taken together, the results with CI and NH listeners suggest that CI users' increased noise susceptibility is primarily a result of reduced spectral resolution and the loss of fine spectro-temporal cues. Spectral cues transmitted by the limited number of implanted electrodes are further reduced by the degree of channel interaction between electrodes. Cochlear implant users are unable to access the important spectro-temporal cues needed to separate speech from noise, including fluctuating noise, because of severely reduced functional spectral resolution. Improved spectral resolution and better channel selectivity may allow CI users

to better understand speech in challenging, noisy listening conditions.

ACKNOWLEDGMENTS

We wish to thank the cochlear implant listeners and normal-hearing subjects for their participation in the current experiments. We also would like to thank John J. Galvin III for his assistance in editing the manuscript. This research was supported by NIDCD 1DC004993.

REFERENCES

- BACON SP, OPIE JM, MONTOYA DY. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J. Speech Hear. Res.* 41:549–563, 1998.
- BAER T, MOORE BCJ. Effects of spectral smearing on the intelligibility of sentences in noise. *J. Acoust. Soc. Am.* 94:1229–1241, 1993.
- BAER T, MOORE BCJ. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J. Acoust. Soc. Am.* 95:2277–2280, 1994.
- BOOTHROYD A, MULHEARN B, GONG J, OSTROFF J. Effects of spectral smearing on phoneme and word recognition. *J. Acoust. Soc. Am.* 100:1807–1818, 1996.
- DOWELL RC, SELIGMAN PM, BLAMEY PJ, CLARK GM. Speech perception using a two-formant 22-electrode cochlear prosthesis in quiet and in noise. *Acta Otolaryngol. (Stockh.)* 104:439–446, 1987.
- EISENBERG LS, DIRKS DD, BELL TS. Speech recognition in amplitude modulated noise of listeners with normal and impaired hearing. *J. Speech Hear. Res.* 38:222–233, 1995.
- FESTEN JM, PLOMP R. Effects of fluctuation noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88:1725–1736, 1990.
- FU QJ, SHANNON RV AND WANG, X. Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *J. Acoust. Soc. Am.* 104: 3586–3596, 1998.
- GUSTAFSSON HA, ARLINGER SD. Masking of speech by amplitude-modulated noise. *J. Acoust. Soc. Am.* 95:518–529, 1994.
- HALL JW, GROSE JH. Some effects of auditory grouping factors on modulation detection interference (MDI). *J. Acoust. Soc. Am.* 90:3028–3035, 1991.
- HOCHBERG I, BOOTHROYD A, WEISS M, HELLMAN S. Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear Hear.* 13:263–271, 1992.
- HORST JW. Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects. *J. Acoust. Soc. Am.* 82:874–885, 1987.
- KIEFER J, MÜLLER J, PFENNIGDORF T, SCHÖN F, HELMS J, VON ILBERG C, GSTÖTTNER W, EHRENBERGER K, ARNOLD W, STEPHAN K, THUMFART S, BAUR S. Speech understanding in quiet and in noise with the CIS speech coding strategy (Med-El Combi-40) compared to the Multipeak and Spectral peak strategies (Nucleus). *ORL* 58:127–135, 1996.
- MÜLLER-DEILER J, SCHMIDT BJ, RUDERT H. Effects of noise on speech discrimination in cochlear implant patients. *Ann. Otol. Rhinol. Laryngol.* 166:303–306, 1995.
- NEJIME Y, MOORE BCJ. 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, 1997.

- NELSON PB, JIN S-H. Factors affecting speech understanding in gated interference: cochlear implant users and normal-hearing listeners. *J Acoust. Soc. Am.* 115:2286–2294, 2004.
- NELSON PB, JIN S-H, CARNEY AE, NELSON DA. Understanding speech in modulated interference: cochlear implant users and normal-hearing listeners. *J. Acoust. Soc. Am.* 113:961–968, 2003.
- NILSSON MJ, SOLI SD, SULLIVAN J. Development of a hearing in noise test for the measurement of speech reception threshold. *J. Acoust. Soc. Am.* 95:1985–1999, 1994.
- PLOMP R. A signal-to-noise ratio model for the speech reception threshold of the hearing impaired. *J. Speech Hear. Res.* 29:146–154, 1986.
- QIN MK, OXENHAM AJ. Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.* 114:446–454, 2003.
- SKINNER MW, CLARK GM, WHITFORD LA, et. al. Evaluation of a new spectral peak coding strategy for the Nucleus-22 channel cochlear implant system. *Am. J. Otol.* 15 (Suppl. 2):15–27, 1994.
- TAKAHASHI GA, BACON SP. Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear. Res.* 35:1410–1421, 1992.
- TER KEURS M, FESTEN JM, PLOMP R. Effect of spectral envelope smearing on speech reception. I. *J. Acoust. Soc. Am.* 91:2872–2880, 1992.
- TER KEURS M, FESTEN JM, PLOMP R. Effect of spectral envelope smearing on speech reception. II. *J. Acoust. Soc. Am.* 93:1547–1552, 1993.
- YOST WA. Auditory image perception and amplitude modulation: frequency and intensity discrimination of individual components for amplitude-modulated two-tone complexes. In: Cazals L, Demany L, Horner K (Eds) *Auditory Physiology and Perception*. Pergamon, Oxford, pp 487–493, 1992.