Amplitude Mapping and Phoneme Recognition in Cochlear Implant Listeners

Fan-Gang Zeng and John J. Galvin III

Objective: **Speech and other environmental sounds must be compressed to accommodate the small electric dynamic range in cochlear implant listeners. The objective of this paper is to study whether and how amplitude compression and dynamic range reduction affect phoneme recognition in quiet and in noise for cochlear implant listeners.**

Design: **Four implant listeners using the Nucleus-22 SPEAK speech processor participated in this study. The amount of compression was varied by manipulating the Q-value in the SPEAK processor. The size of the dynamic range was systematically reduced by increasing the threshold level and decreasing the comfortable level in the processor. Both female- and male-talker vowel and consonant materials were used to evaluate speech recognition performance in quiet and in noise. Speech-spectrum-shaped noise was mixed with the speech signal and presented continuously to the speech processor through a direct electric connection. Signal to noise ratios were changed over a 30 to 40 dB range, within which phoneme recognition increased from chance to asymptotic performance. Phoneme recognition scores were obtained as the number of active electrodes was reduced from 20 to 10 to 4. For purposes of comparison, phoneme recognition data also were collected in four normal-hearing listeners under comparable laboratory conditions.**

Results: **In both quiet and noise, the amount of amplitude compression did not significantly affect phoneme recognition. The reduction of dynamic range marginally affected phoneme recognition in quiet, but significantly degraded phoneme recognition in noise. Generally, the 20- and 10-electrode processors produced similar performance, whereas the 4-electrode processor produced significantly poorer performance. Compared with normal-hearing listeners, cochlear-implant listeners required higher signal to noise ratios to achieve comparable recognition performance and produced significantly lower recognition scores at the same signal to noise ratios.**

Conclusions: **The amount of amplitude compression does not significantly affect phoneme recognition, whereas reducing dynamic range significantly lowers phoneme recognition, particularly in noise and for vowels. Because the SPEAK processor extracts mostly spectral peaks, the present conclusions may not be applied to other types of processors extracting temporal envelope cues. The present results also suggest that more than four electrodes are required to optimize speech recognition in multiple-talker and noise conditions. A significant performance gap in speech recognition still remains between cochlear implant and normal-hearing listeners at the same signal to noise ratios. Improved cochlear implant designs and fitting procedures are required to narrow and, hopefully, close this performance gap.**

(Ear & Hearing 1999;20;60–74)

A normal-hearing person can not only detect sound-driven vibrations of only a half-nanometer (the diameter of an atom), but also process acoustic information that varies by at least 10 orders of magnitude in intensity (Hudspeth, 1997). This 100 dB acoustic dynamic range is necessary because environmental sounds, including speech and music, often change in intensity over a 30 to 60 dB range and are presented to a listener at varying overall levels (Boothroyd, Erickson, & Medwetsky, 1994; Fletcher, 1953). This large dynamic range, coupled with fine intensity resolution (200 discriminable steps) and spectral and temporal tuning, allow a normal-hearing listener to maintain high speech intelligibility in noisy backgrounds and at presentation levels from 40 to 110 dB SPL (Borg & Zakrisson, 1973; Viemeister, 1988).

In contrast, a cochlear implant listener typically has a dynamic range of 10 to 20 dB and 20 discriminable steps (Nelson, Schmitz, Donaldson, Viemeister, & Javel, 1996; Zeng, Shannon, & Hellman, 1998). In addition, loudness grows differently in electric stimulation of the auditory nerve than in acoustic stimulation (Zeng & Shannon, 1992). This small dynamic range and abnormal loudness growth are likely due to a combination of factors, including the loss of cochlear compression and abnormal recruitment of nerve activity in electric stimulation (Zeng & Shannon, 1994; Zeng et al., 1998). To accommodate the small electric dynamic range and to restore normal loudness growth, all cochlear implant processors have to compress acoustic amplitudes. However, the type and the amount of ampli-

Department of Hearing and Speech Sciences and Center for Comparative Neuroscience (F.-G. Z.), University of Maryland, College Park, Maryland; and House Ear Institute (J.J.G.), Los Angeles, California.

0196/0202/99/2001-0060/0 • Ear & Hearing •

tude compression remains a difficult problem in fitting a speech processor.

Most cochlear implant speech processors use an automatic gain control circuit or a sensitivity control to compensate for overall level differences due to the talker's vocal effort, distance and position between the talker and the microphone. However, implant speech processors differ vastly in the methods used to compress or map the roughly 30 dB speech dynamic range into the individual implant user's electric dynamic range. Early single-channel cochlear implants such as the 3 M/House device had a step-like amplitude mapping function that clipped essentially all acoustic amplitudes above zero (House, 1978). For multi-electrode speech processors employing a compressed analog strategy (Ineraid device, see Eddington, 1980 and Eddington, Dobelle, Brackmann, Mladejovsky, & Parkin, 1978; USCF/Storz device, see Merzenich, Rebscher, Loeb, Byers, & Schindler, 1984), an automatic gain control was used to compress or restrict the wide acoustic dynamic range down to the narrow electric dynamic range. More recently, speech processors employing the continuous interleaved sampling (CIS) strategy (both Clarion and MED-EL devices) have used a logarithmic mapping function to compress the acoustic dynamic range (Wilson, Finley, Lawson, Wolford, Eddington, & Rabinowitz, 1991). On the other hand, speech processors in the Nucleus device, including the SPEAK processor, map the acoustic-to-electric amplitude using a power function with variable exponents to control the amount of compression (Cochlear Corporation, 1995, pp. 38–41). Although one desirable goal of amplitude mapping in cochlear implants is to restore normal loudness growth (Zeng & Shannon, 1992, 1994), no systematic benefits in intelligibility have been achieved through the loudness-matching amplitude mapping strategy (Boex, Eddington, Noel, Rabinowitz, Tierney, & Whearty, Reference Note 1). At present, no standard exists for the appropriate mapping function from the acoustic amplitude to the electric amplitude, and little is known about the perceptual effects of these various amplitude mapping functions.

This study examines to what extent the amplitude mapping parameters that are available in the Nucleus SPEAK processor affect speech recognition in quiet and in noise. The first experiment manipulated the amount of compression in the acoustic-to-electric amplitude mapping function and the second experiment manipulated the size of electric dynamic range. These amplitude mapping parameters were studied as a function of the number of active electrodes in both quiet and noise conditions. For purposes of comparison, normal-hearing listeners were also tested at identical speech-to-noise ratios under similar laboratory conditions.

METHODS

Participants and Their Processor Parameters

Four Nucleus-22™ cochlear implant users using the SPEAK strategy participated in this study. Table 1 shows the age, gender, deafness etiology, speech recognition results and other processor information. These participants, based on their sentence and word recognition scores, were average to excellent users of the cochlear implants (Skinner et al., 1994). No poor users were chosen in this study to avoid a floor effect. All implant participants had extensive experience in various psychophysical and speech experiments.

The Nucleus processor employing the SPEAK strategy divides an input acoustic signal into 20 frequency bands, extracts the amplitude envelope from all 20 bands, and stimulates the electrodes corresponding to the 6 to 10 bands with the maximal amplitude (McDermott, McKay, & Vandali, 1992). The SPEAK strategy converts a 30 dB acoustic range (between a base level 4 and a maximal level 150 in linear digitized amplitude) into the electric dynamic range between the threshold (T-level) and the maximal comfortable level (C-level). The SPEAK processor also allows different preamplifier gain (G) for each electrode. All participants used a default gain setting of 8 (except for participant RK, who used a gain setting of 6 on active electrodes 1 and 3). For the 4-, 10-, and 20-electrode experimental processors, T-levels and C-levels were directly measured, and all electrodes were loudness-balanced at the C-level with an electrode sweep method (Co-

TABLE 1. Biographical and audiological information for cochlear implant (CI) participants in this study.

Subject	Age	Gender	Etiology	Years as CI User	Frequency Table	Stimulation Mode	CUNY Sentence Score % Correct	NU6 Word Score % Correct
DJ	55	F	Hereditary	8	9	BP + 1	100	60
EB	56	M	Trauma/Unknown	8	7	BP + 1	79	24
JM	40	M	Trauma	6	9	BP + 1	99	70
RK	55	M	Unknown	3	9	BP + 1	99	48

chlear Corporation, 1995, p. 67). The 4- and 10-electrode processors were created in a fashion similar to the Fishman, Shannon, and Slattery (1997) study: for the 4-electrode processor, five adjacent frequency bands were summed and mapped to four equally spaced electrode pairs, and for the 10-electrode processor, two adjacent frequency bands were summed and mapped to 10 equally spaced electrode pairs. Figure 1 plots the T-level and the C-level as a function of active electrode position for the 20-electrode processor. The T-level and C-level are plotted in terms of either clinical units (right y-axis) or physical units (left y-axis, dB re:1nC charge). Based on each subject's amplitude calibration table obtained from Cochlear Corporation, clinical unit 1 was calculated to be equal to 13.98, 15.56, 15.56, and 14.81 dB for participants DJ, EB, JM, and RK, respectively; similarly, the maximal clinical unit 239 was 52.44, 51.49, 51.71, and 51.73 dB for participants DJ, EB, JM, and RK, respectively. If the clinical units were logarithmic, then they would produce dynamic ranges identical to that represented by the charge units in dB. However, Figure 1 shows that the dynamic range represented by physical units is slightly smaller than that represented by clinical units. The average dynamic range across all electrodes was 13.46, 6.01, 8.50, and 11.07 dB for participants DJ, EB, JM, and RK, respectively. These dynamic ranges were comparable with that in the Skinner, Holden, Holden, Demorest, and Fourakis (1997) study. For the 4- and 10-electrode experimental processors, the T- and C-levels were re-measured and found to be similar to the values obtained with the 20-electrode processor. Specifically, the mean differences in clinical units between the 20- and 4-electrode processors, averaged across all subjects and the four common electrodes, were 1.6 (range = 0 to 10) for the T-level and 2.7 (range = 0 to 15) for the C-level, representing only 1.4% and 1.9% increase from the baseline T- and C-levels in the 20-electrode processor. These similar T- and C-levels suggested that loudness summation due to nonsimultaneous stimulation across different electrodes was not a significant factor in the present implant participants using the SPEAK strategy.

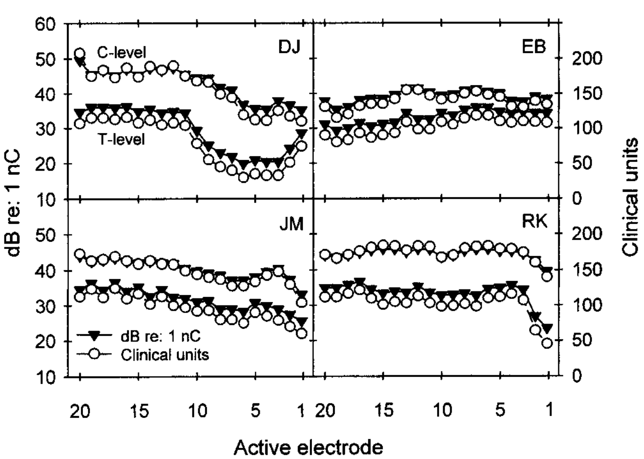

Figure 1. Threshold (T-level) and maximal comfortable loudness (C-level) data represented in either clinical units (open circles) or electric charge (filled triangles) for each subject's clinically assigned processor. The x-axis represents active electrode number, the left y-axis represents electric charge (dB re: 1 nC charge), and the right y-axis represents clinical units.

For purposes of comparison, four male normal-hearing listeners, aged 27–35 yr, also participated in the vowel and consonant recognition experiments. These listeners all had normal hearing, with 20 dB HL or below for octave frequencies between 250 and 8000 Hz. Both implant and normal-hearing listeners were paid for their participation in this study and received formal informed consent.

Stimuli

Stimuli were medial vowels and consonants, spoken by one male and one female talker. The 12 vowels included: ɑ, æ, ʌ, ɔ, ɛ, ɜ, e, I, i, o, U, and u in h/V/d format (Hillenbrand, Getty, Clark, & Wheeler, 1995). The 16 consonants included: b, d, g, p, t, k, f, θ, s, ʃ, v, ɚ, z, ʒ, m, and n in a/C/a format (Turner, Souza, & Forget, 1995). The Hillenbrand vowels were 16-bit .WAV files and sampled at 16 kHz, and the Turner consonants were 16-bit .WAV files and sampled at 44.1kHz. These vowel and consonant stimuli were output via a PC soundcard (Turtle Beach MultiSound Fiji board) connected to one channel of a mixer (Tucker-Davis Technologies, TDT SM1).

The speech-spectrum-shaped noise was generated by passing a white noise (TDT WG1) through a specially designed low-pass filter with a cut-off frequency at 608 Hz and a −12 dB/octave slope (Byrne et al., 1994). The noise then was attenuated (TDT PA1) to achieve the average signal to noise ratios of −20, −15, −10, −5, 0, 10, 20 dB. The 40 dB signal to noise ratio range would assure a full range of speech recognition scores from chance to plateau levels (Hochberg, Boothroyd, Weiss, & Hellman, 1992). The noise was delivered to another channel of the mixer where it was summed with the phonemic stimuli.

For normal-hearing listeners, the summed speech and noise stimuli were amplified (Crown D-75) and presented at a comfortably loud level (about 75 dBA) monaurally via an insert earphone (Etymotic Research ER 2) in a double-walled sound-treated booth (IAC). For cochlear implant listeners, the mixed stimuli were presented directly from the output of the mixer to the external input of the Nucleus Spectra-22 processors via the Audio Input Selector (AIS). The AIS not only provided electrical isolation but also served as a linear amplifier or attenuator. The gain of the AIS was always set at 4 in the present study. The speech processor was set to the normal (N) mode. The sensitivity control was adjusted by the listener to achieve comfortable listening level while acclimatizing to each experimental processor; once the sensitivity was set, it was not moved during the entire testing session. In practice, the sensitivity control was set at or near 2.5. The direct connection to the speech processor, bypassing the processor's microphone, clearly did not represent a realistic everyday listening experience, but did allow isolation of the amplitude-mapping factor in speech recognition from other factors (such as reverberation, ambient noise floor, and acoustic-electrical properties of microphone and speaker) that are present in the speaker-microphone connection.

Acoustic-to-Electric Amplitude Mapping

Two manipulations were examined in the acoustic-to-electric amplitude mapping function. The first manipulation systematically decreased the amount of compression by varying the Q-value from 20 to 50 in 10 Q-value steps (Fig. 2a). According to the Cochlear Corporation *Technical Reference Manual* (p. 41), the Q value represents the percent reduction from the C-level at the 50 digital acoustic amplitude, or 20 dB above the threshold level, as indicated on the y-axis in Figure 2a. A power function was fitted to the curves in Figure 2a and the exponent of the power function was estimated to be 0.24, 0.40, 0.51, and 0.63 for Q values of 20, 30, 40, and 50, respectively. In this experiment, the electric dynamic range was unchanged from the original T- and C-levels in the clinically fitted processor. The second manipulation systematically decreased the original dynamic range to 25% by increasing the T-levels to 75% of the dynamic range, and to a binary representation of the acoustic amplitude by additionally decreasing the C-levels to 76% of the dynamic range (see Fig. 2b). In the reduced electric dynamic range experiment, the Q-value was unchanged from the original value (20 except for RK whose Q value equals 31) in the clinically fitted processor. The acoustic level where the electric amplitude of stimulation was switched to 75% of the dynamic range was about 19 dB and 24 dB above the acoustic threshold (i.e., base level = 4) for Q-20 and Q-30 conditions, respectively.

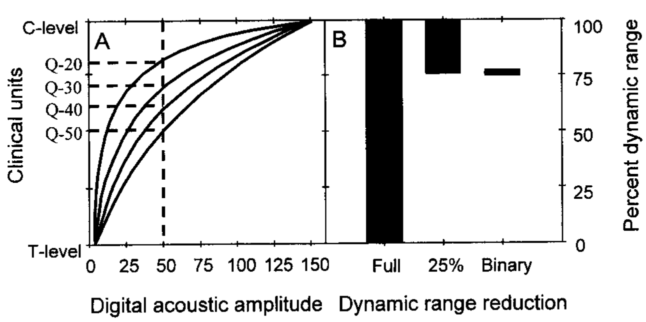

Figure 2. Amplitude compression (panel A) and dynamic range reduction (panel B). In panel A, the different compression ratios are represented by different Q values depicted along each curve. The Q value is defined as the percent drop from the C-level when the digital amplitude is 10 dB from maximum. In panel B, the dynamic range is reduced to 25% by increasing the T-level to 75% of the dynamic range, and to a binary representation of acoustic amplitudes by a combination of increasing the T-level to 75% of the dynamic range and decreasing the C-level to 76% of the dynamic range.

Because spectral information is reduced when a smaller number of electrodes was activated, amplitude compression and dynamic range reduction might have a more significant effect on speech recognition because the implant user would be forced to rely more on temporal envelope cues. The effects of these amplitude mapping manipulations on speech recognition were therefore studied as a function of the number of electrode (4, 10, and 20). Together, this study generated a total of 18 experimental processors for each cochlear implant listener (3 electrode conditions × 4 Q-value conditions + 2 dynamic range conditions)]. For speech recognition in noise, only three representative conditions were tested, including: 1) Q-20 with the full dynamic range; 2) Q-50 with the full dynamic range; and 3) Q-20 with binary representation of the dynamic range. These three amplitude conditions were tested under six different signal to noise ratios including the quiet condition, resulting in a total of 18 additional test sessions each for vowels and consonants.

Figure 3 shows a spectrogram (top panel) and three representative electrodograms (bottom three panels) for the speech stimulus /ʌsʌ/. Note in the second panel that both the spectral and temporal characteristics of the stimulus /ʌsʌ/ are well preserved in the electrodogram of the 20-electrode processor (Q = 20). Note in the third panel that the binary representation of the acoustic amplitudes also largely preserved the spectral distribution of the original spectrogram but only minimally preserved the temporal envelope cues. Finally, note in the bottom panel that most spectral details are lost in the electrodogram of the 4-electrode processor.

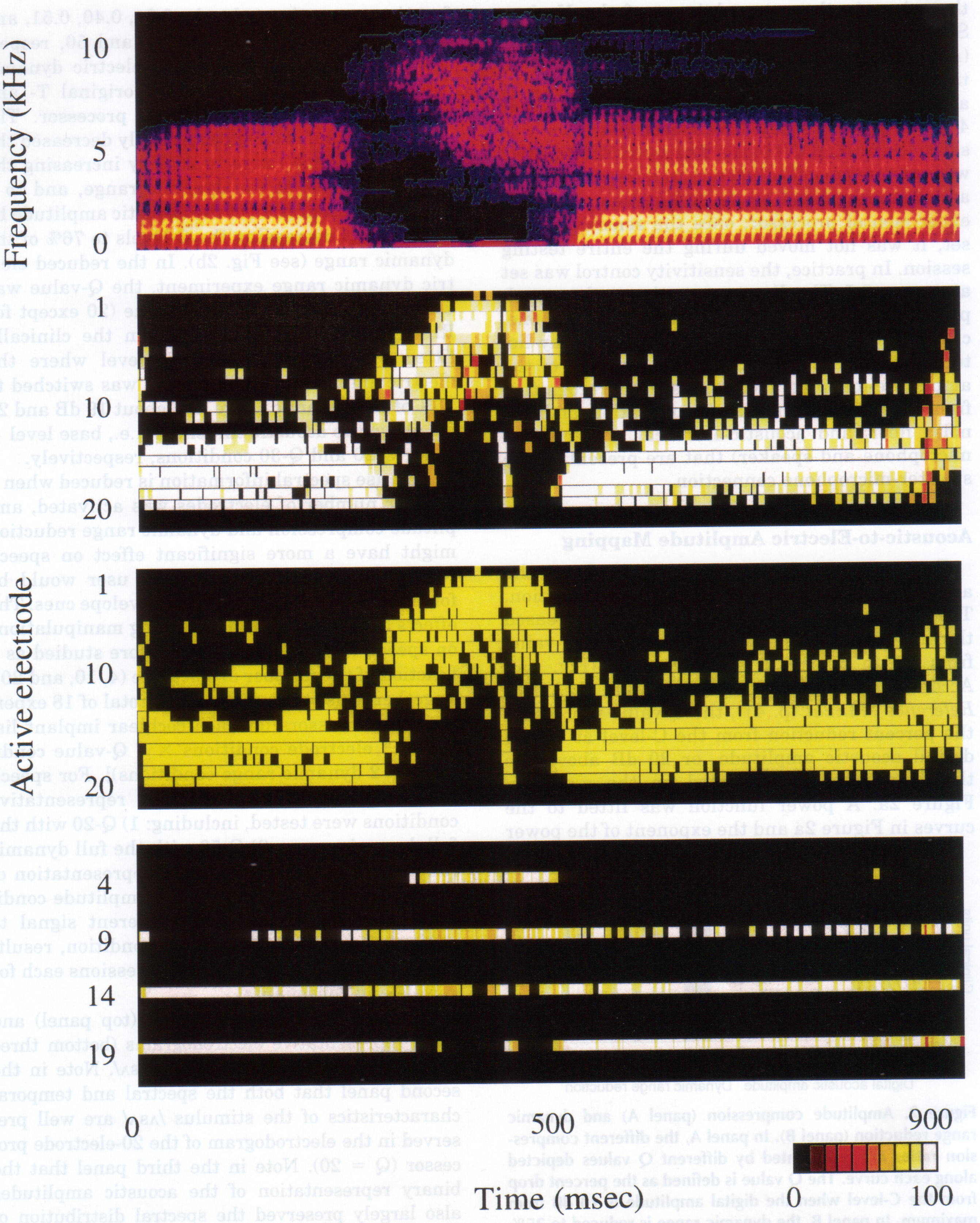

Figure 3. Spectrogram (top panel) and electrodograms of the phoneme /ʌsʌ/ are shown for subject JM for the following experimental processor conditions: 20-electrode and Q-20 (second panel), 20-electrode, Q-20, and binary dynamic range (third panel), and 4-electrode and Q-50 (bottom panel). For each panel, the x-axis represents time (msec), the y-axis represents active electrode from most apical (20) to most basal (1), and the stimulation level is referenced to the color scale.

Procedures

Vowel and consonant recognition were conducted separately in a closed-set format using an interface developed at the House Ear Institute. The test order of these experimental conditions was: amplitude compression, dynamic range reduction, and phoneme recognition in noise. Within each condition, the test order was pseudo-randomized for all listeners across different compression ratios, dynamic ranges, and signal to noise ratios. All listeners were given 15 minutes to acclimate to each experimental processor and were allowed to preview all stimuli before formal test sessions. Each test session consisted of five presentations for each phoneme by each of the two talkers. The order of each phoneme's occurrence in each test session was randomized. The listener's response to the speech stimulus was stored as a confusion matrix. No trial-by-trial feedback was given regarding the correctness of the response.

RESULTS

Amplitude Compression

Because all individuals showed a similar pattern of results, only the group mean data are reported.

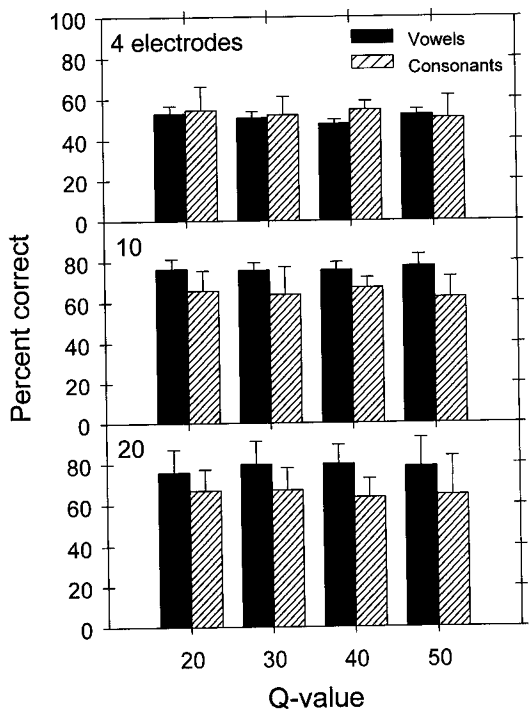

Figure 4. Averaged vowel and consonant recognition scores as a function of electrode number and amount of amplitude compression. The top, middle, and bottom panels show performance for the 4-, 10-, and 20-electrode processor, respectively. In each panel, the x-axis represents the Q-value, from most-compressive (Q-20) to least-compressive (Q-50). The filled bars represent vowel scores and the shaded bars represent consonant scores. Error bars represent 1 SD.

Figure 4 presents the averaged data for the vowel (filled bars) and consonant (shaded bars) recognition as a function of amplitude compression (x-axis) and as a function of the number of electrodes (rows). No systematic effect of amplitude compression was noted in both vowel and consonant recognition. A 2-way analysis of variance (ANOVA) confirmed that there was no significant difference between the compression conditions [vowels: $F(3, 36) = 0.02$; $p = 0.99$; consonants: $F(3, 36) = 0.11$; $p = 0.96$], but a significant effect for the number of electrodes [vowels: $F(2, 36) = 61.57$; $p < 0.01$; consonants: $F(2, 36) = 5.18$; $p = 0.01$]. No significant interactions were found between compression and electrode number factors. A post hoc Scheffe test on vowel recognition indicated no significant difference between the 10- and the 20-electrode processor ($p > 0.05$), but significantly poorer performance for the 4-electrode processor ($p < 0.01$). The same test on consonant recognition also revealed significantly better performance for the 20-electrode processor than the 10- and 4-electrode processors ($p < 0.01$).

To demonstrate the electrode number effect more clearly, a grand average was performed across all participants and all Q-value conditions. Figure 5 shows the averaged vowel and consonant recognition as a function of the number of electrodes. For vowel recognition, the percent correct score decreased from 78% with 20 electrodes to 73% with 10 electrodes to 49% with four electrodes; for consonant recognition, the percent correct score decreased from 66% with 20 electrodes to 59% with 10 electrodes to 53% with four electrodes. The abrupt change in vowel recognition versus the gradual change in

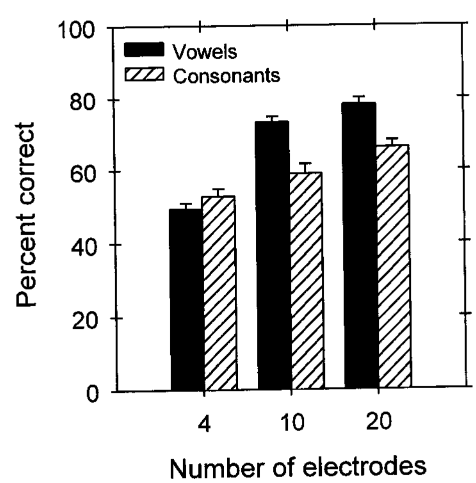

Figure 5. Averaged scores across all subjects and Q-conditions are shown as a function of the number of active electrodes. The x-axis represents the number of electrodes, the y-axis represents percent correct. The filled bars represent vowel scores, and the shaded bars represent consonant scores. Error bars represent 1 SD.

consonant recognition as a function of the number of electrodes may be due to the fact that 1) vowel recognition requires finer spectral resolution than consonant recognition, and 2) consonant recognition depends more on nonspectral cues that are not degraded by reducing the number of electrodes (Svirsky & Meyer, Reference Note 4).

Dynamic Range Reduction

In a similar fashion, Figure 6 presents the averaged vowel and consonant recognition data as a function of dynamic range reduction and as a function of the number of electrodes. These data show that, at least under quiet conditions, vowel and consonant recognition was not greatly affected by reducing the electric dynamic range, even if it was reduced to a binary representation of the acoustic amplitude. A 2-way ANOVA indicated that vowel recognition was marginally affected by the dynamic range reduction [$F(2, 27) = 3.71$; $p = 0.04$], whereas consonant recognition was not significantly affected [$F(2, 27) = 0.48$; $p = 0.63$]. Similarly, reducing the number of electrodes significantly decreased vowel recognition [$F(2, 27) = 35.66$; $p < 0.01$] but not consonant recognition [$F(2, 27) = 2.05$; $p = 0.15$].

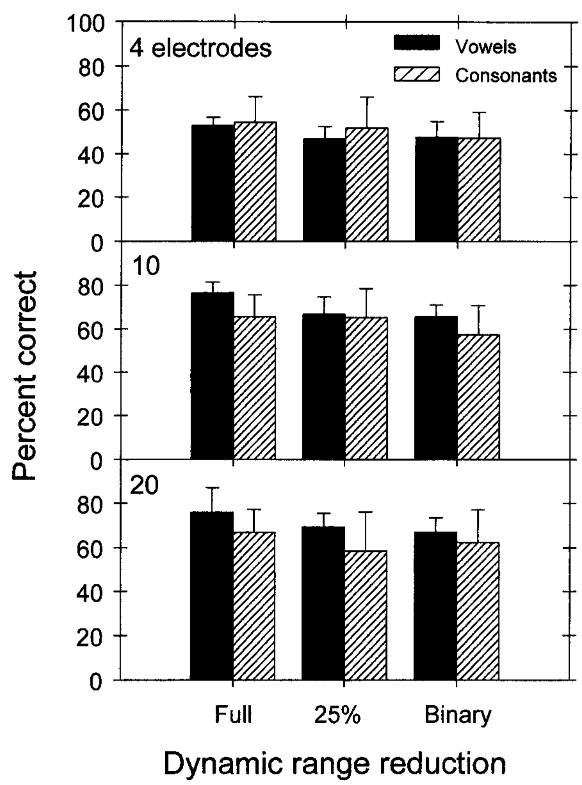

Figure 6. Averaged vowel and consonant recognition scores as a function of electrode number and dynamic range. All symbols are the same as in Fig. 4 except for the x-axis, which represents the manipulation of the electrical dynamic range from full to 25% to binary representation.

Phoneme Recognition in Noise

The previous two experiments demonstrated that amplitude compression and dynamic range reduction generally had little effect on vowel and consonant recognition under quiet conditions. However, under more realistic listening conditions where background noise is always present, different amplitude compression ratios could significantly change the normal intensity ratios between speech sounds and between speech and noise (Hickson & Byrne, 1997; Van Tasell & Trine, 1996). In the noise experiment, recognition in quiet was measured again on the same day when other recognition data in noise were collected, and served as a test-retest reliability measure when compared with the quiet data in the previous two experiments.

Figure 7 shows vowel recognition and Figure 8 shows consonant recognition as a function of signal to noise ratios (x-axis). For normal-hearing listeners, the range of performance (mean and ± standard deviation recognition scores) is shown as a shaded area. The average performance for these normal-hearing listeners rose above the chance level at about −15 dB signal to noise ratio for both vowels and consonants, and reached 50% percent correct level at about −10 dB for vowels and −8 dB for consonants. The vowel recognition reached a 95% asymptotic performance level at about −5 dB signal to noise ratio whereas consonant recognition did not reach the same asymptotic level until the +5 dB ratio. These differences most likely reflect the intensity difference between vowels and consonants (Boothroyd et al., 1994; Fletcher, 1953) and listeners' different abilities to use formant transition and burst cues to recognize the consonants (Zeng & Turner, 1990).

Figures 7 and 8 also show cochlear implant listeners' individual (top four rows) and averaged (bottom row) performance under similar conditions. The amplitude manipulations are represented as different symbols in each panel and the electrode numbers are shown by different columns. In the remeasured quiet conditions (including Q-20-full dynamic range, Q-50-full dynamic range, and Q-20-binary dynamic range), the implant listeners produced significantly better performance than the previous experiments for vowel recognition [71.9% versus 65.3%, paired-t(df = 23) = −4.10; $p < 0.01$], but not for consonant recognition [60.3% versus 61.0%, paired-t(df = 23) = −0.30; $p > 0.10$]. The better vowel recognition in the re-tested condition likely reflected a learning effect. Because the relative difference in vowel recognition was still minimal between the Q-20 and Q-50 conditions (rightmost data points in Fig. 7), this improvement in the

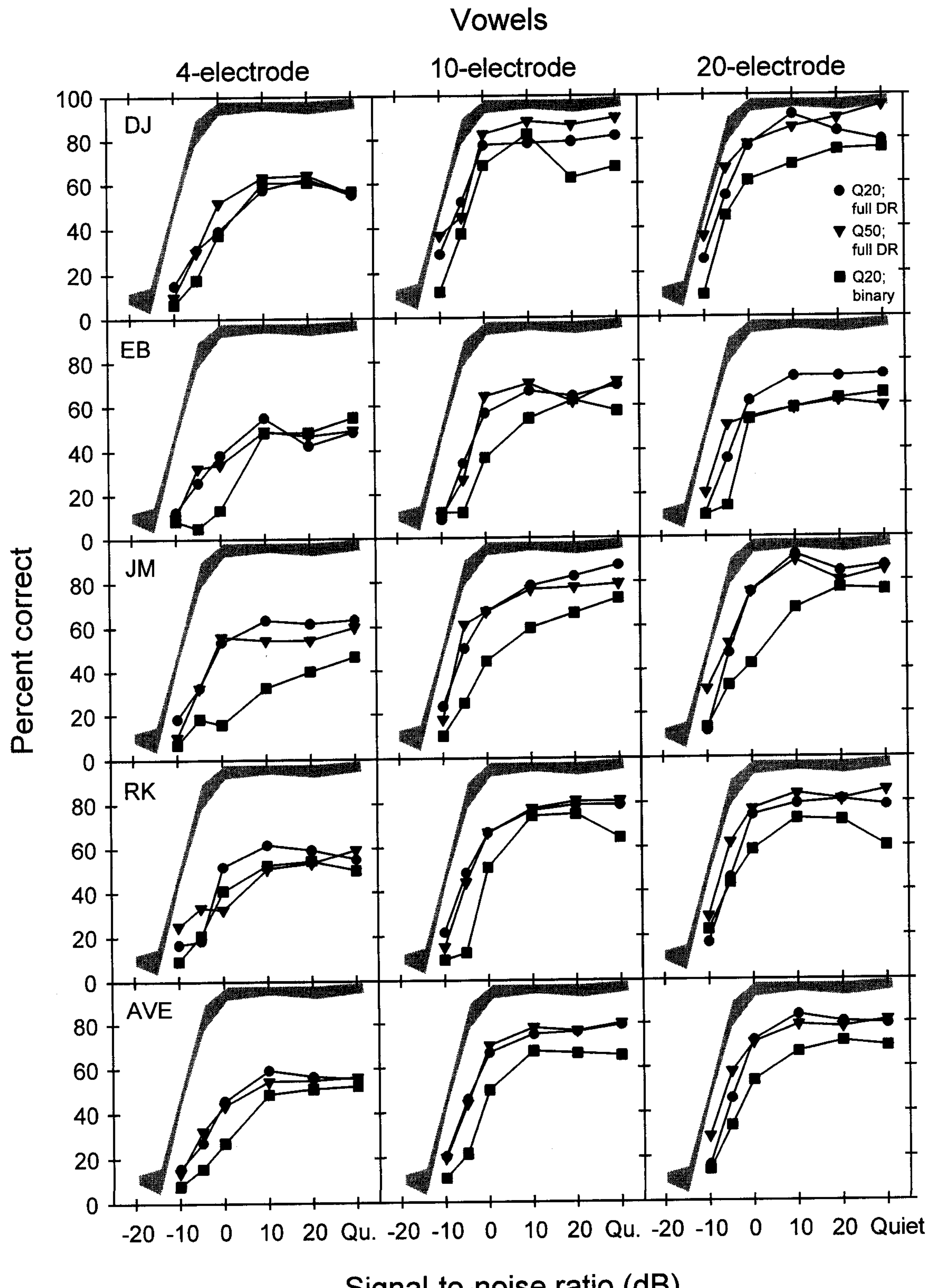

Figure 7. Vowel recognition in noise for normal-hearing and cochlear implant listeners. Three columns, from left to right, show performance for the 4-, 10-, and 20-electrode processor conditions. For each column, the x-axis represents the signal to noise ratio, from −20 dB to +20 dB and in quiet. Five rows, from top to bottom, show performance for DJ, EB, JM, RK, and across-subject average. For each row, the y-axis represents percent correct. For all plots, the circles represent scores in the Q-20 and full dynamic range condition, the inverted triangles represent scores in the Q-50 and full dynamic range condition, and the squares represent scores in the Q-20 and binary dynamic range condition. The shaded area in each plot represents the range of performance for the four normal-hearing listeners (mean and ±standard deviation).

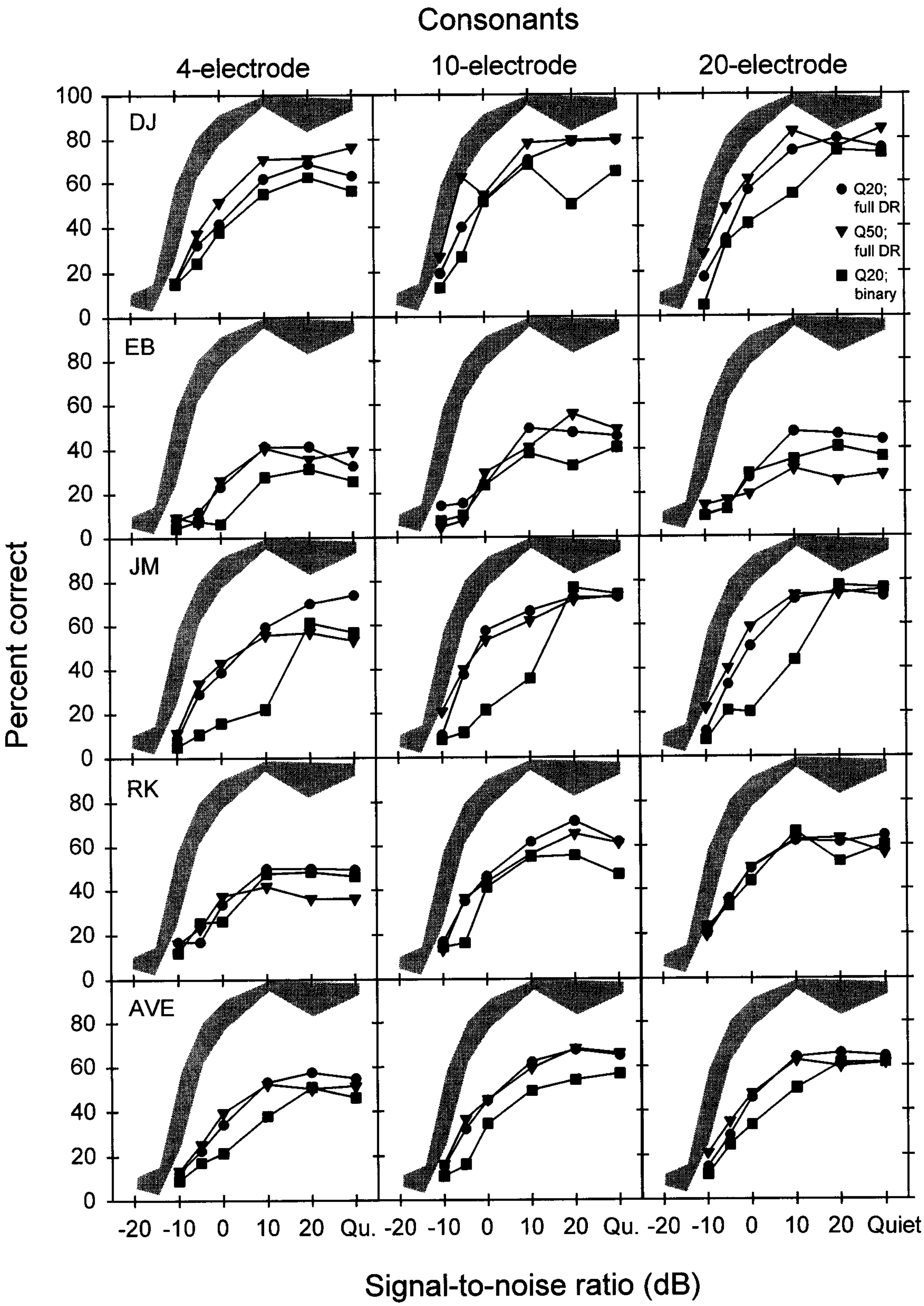

Figure 8. Consonant recognition in noise for normal-hearing and cochlear implant listeners. Three columns, from left to right, show performance for the 4-, 10-, and 20-electrode processor conditions. For each column, the x-axis represents the signal to noise ratio, from −20 dB to +20 dB and in quiet. Five rows, from top to bottom, show performance for DJ, EB, JM, RK, and across-subject average. For each row, the y-axis represents percent correct. For all plots, the circles represent scores in the Q-20 and full dynamic range condition, the inverted triangles represent scores in the Q-50 and full dynamic range condition, and the squares represent scores in the Q-20 and binary dynamic range condition. The shaded area in each plot represents the range of performance for the four normal-hearing listeners (mean and ±standard deviation).

absolute scores would not change the conclusion reached in the earlier experiment.

In the noise condition, except for a few data points (e.g., 20-electrode processor with Q-50 for DJ), cochlear implant listeners performed significantly poorer than the normal-hearing listeners. Their poorer-than-normal performance can be viewed from two perspectives. First, cochlear implant listeners produced much lower recognition scores at the same signal to noise ratio (e.g., one can find the exact difference in performance between normal-hearing and implant listeners for each experimental condition by simply drawing a vertical line at a specific signal to noise ratio). Second, cochlear implant listeners required much higher signal to noise ratios to achieve the same level of recognition performance (e.g., one can find the exact difference in signal to noise ratios by drawing a horizontal line at a specific recognition level that cross both normal-hearing and implant data). We shall return to this point in the discussion section.

A 3-way ANOVA (amplitude mapping, the number of electrodes, and signal to noise ratios) was performed to examine quantitatively the effects of amplitude mapping and electrode number on vowel and consonant recognition in noise. Similar to the quiet condition, electrode number produced a highly significant effect on both vowel and consonant recognition in noise [$F(2, 162) = 109.02$; $p < 0.01$ for vowels and $F(2, 162) = 9.37$; $p < 0.01$ for consonants]. A post hoc Scheffe test revealed that for vowel recognition, the 10-electrode and 20-electrode processors did not produce significantly different performance ($p = 0.62$); however, both were significantly better than the 4-electrode processor ($p < 0.01$). For consonant recognition, the 10- and 20-electrode processors were significantly better than the 4-electrode processor ($p = 0.05$).

Amplitude mapping also produced a significant effect on both vowel and consonant recognition in noise [$F(2, 162) = 45.53$; $p < 0.01$ for vowels and $F(2, 162) = 10.00$; $p < 0.01$ for consonants]. A post hoc Scheffe test revealed that, although the Q-20 and Q-50 conditions produced no significantly different performance ($p > 0.5$), both were significantly better than the binary dynamic range condition ($p = 0.01$ for vowel recognition; $p = 0.05$ for consonant recognition). However, this reduction in performance for the binary condition was relatively small for all implant participants except for JM. No interactions among all variables were found ($p > 0.10$), except for a significant interaction between the number of electrodes and the signal to noise ratio in vowel recognition ($p < 0.01$). Vowel recognition with a 4-electrode processor was further degraded at low signal to noise ratios, indicating that more than four electrodes are needed under more adverse listening conditions to preserve the spectral details of the speech signals.

DISCUSSION

Under the present laboratory conditions and for the Nucleus SPEAK processor, no significant effect of amplitude compression was found on vowel and consonant recognition in both quiet and noise, but reducing dynamic range to a binary representation of the acoustic amplitude significantly degraded phoneme recognition, particularly in noise and for vowels. Compared with the normal-hearing listeners under noise conditions, cochlear implant listeners (including the "star implant users") produced significantly poorer phoneme recognition performance. The following discussion addresses issues of laboratory versus real-world conditions, relations of the present study to previous studies, and the practical meaning and theoretical significance of the present study.

Laboratory Conditions

The purpose of this study was to examine whether and how amplitude compression and dynamic range reduction can affect significantly speech recognition under laboratory conditions. For this reason, a direct electric connection from the sound card output to the implant processor was used. The direct connection minimized both the background noise floor and spectral smearing due to reverberation, and avoided any signal irregularities due to microphone differences. In everyday listening conditions, the speech presentation level may be soft and may not be optimally amplified. For example, Skinner, Holden, Holden, Demorest, and Fourakis (1997) and Skinner, Holden, and Holden (1997) showed that overall speech recognition performance generally decreased when speech presentation level was reduced from a raised-to-loud vocal effort level of 70 dB SPL to a soft level of 50 dB SPL. However, many cochlear implant listeners will compensate for subthreshold audibility problems by simply increasing the sensitivity control setting. Increasing the sensitivity control setting can help amplify soft speech levels, but it will also amplify the ambient noise found in everyday listening conditions. As discussed later, speech recognition under lower signal to noise ratios still presents a significant problem for cochlear implant users.

Amplitude Compression

We did not find significant differences in either vowel or consonant recognition as a function of amplitude compression (Q value changed from 20 to 50). We also analyzed percentages of information transferred for voicing, manner, and place features in consonant recognition and found no distinctive effects of amplitude compression on specific features. These results were not really surprising when considering previous results on compression obtained in normal-hearing listeners and hearing aid users. In normal-hearing listeners, compression on temporal envelopes was found to have no effect on speech recognition (Souza & Turner, 1996; Van Tasell & Trine, 1996). In hearing-impaired listeners, similar results were found by either compressing the temporal envelope (Souza & Turner, 1996) or changing the relative consonant-vowel amplitude ratios (Hickson & Byrne, 1997). In similar laboratory conditions, previous studies found no significant compression effect on phoneme recognition until much more severe compression than the present range was used (Cosendai & Pelizzone, Reference Note 2; Fu & Shannon, 1998; Shannon, Zeng, & Wygonski, 1992).

A more compressive amplitude mapping function can be achieved by turning on the noise suppression (S) mode in the Nucleus Spectra processor, which essentially clips lower sound amplitudes and provides more compressive mapping than the normal (N) mode at low acoustic amplitude (Cochlear Corporation, 1995, p. 126). This noise suppression setting has been shown to improve speech recognition in noise but degrade speech recognition in quiet (Muller-Deile, Schmidt, & Rudert, 1995). We also compared the noise suppression setting and the normal setting in one implant listener and found similar results. In quiet, the vowel and consonant recognition scores were 79% and 66% for the normal mode, and were 79% and 47% for the suppression mode; in the 0 dB signal to noise ratio condition, the vowel and consonant recognition scores were 58% and 38% for the normal mode, and were 63% and 44% for the suppression mode. This finding suggests that more compressive amplitude mapping than was used in the present study may produce significant reductions in performance at high signal to noise ratios and significant improvements in performance at low signal to noise ratios.

Next, we noted a small but consistent cross-over pattern between psychometric functions of the Q-20 and the Q-50 conditions for both vowel and consonant recognition (bottom-right panels of Figs. 7 and 8, and more clearly shown by the fitted functions in Fig. 9). The more compressive Q-20 condition produced poorer performance than the Q-50 condition at low signal to noise ratios but better performance at high signal to noise ratios. This cross-over pattern can also be reflected by the steeper slope of the psychometric function for the Q-20 condition than the Q-50 condition. More systematic studies with more sensitive test materials (e.g., words or sentences) are needed to determine how different degrees of amplitude compression affect speech recognition across a wide range of signal to noise ratio conditions.

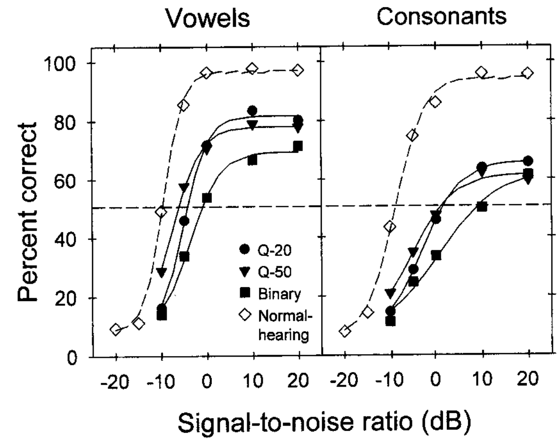

Figure 9. Vowel (left panel) and consonant (right panel) recognition data in noise for averaged normal-hearing and implant listeners. The implant listeners' scores were averaged from performance using the 20-electrode processor. The circles represent scores in the Q-20 and full dynamic range condition, the inverted triangles represent scores in the Q-50 and full dynamic range condition, the squares represent scores in the Q-20 and binary representation condition, and the open diamonds represent scores for normal-hearing listeners. A logistic function has been fit to each set of cochlear implant listeners' data (solid lines) and normal-hearing listeners' data (dashed line). The horizontal dashed line represents the 50% absolute recognition score.

In addition, the input acoustic stimuli can also play a significant role in evaluating the amplitude mapping function in auditory prostheses. Lippmann, Braida, and Durlach (1981) compared multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss and found that amplitude compression was superior to linear amplification only when speech materials with significant level variations were used and when the input speech level was reduced. If the results in the Lippman et al. study can be extended to cochlear implants, then we would expect a more significant effect of amplitude mapping under conditions where more talkers were used and their overall vocal levels were not normalized. These stimulus manipulations are important because they represent real-life communication situations and should be evaluated in future investigations.

Finally, amplitude compression may significantly affect other aspects of speech perception such as perceived sound quality. We did not quantitatively study the effect of amplitude compression on subjective sound quality, but were frequently reminded by our implant listeners that different Q values produced noticeably different sound qualities.

Encoding Loudness in Cochlear Implants

Loudness-related factors may also contribute to the present observation of minimal effects of amplitude compression on phoneme recognition. One factor may be due to the use of electric charge to encode loudness in the Nucleus implant, in which electric pulse amplitude is linearly traded for electric pulse duration (Cochlear Corporation, 1995, p. 37). Zeng, Galvin, and Zhang (1998) systematically measured equal loudness curves as a function of pulse amplitude and duration at both threshold and suprathreshold levels as well as loudness balance functions between pulse amplitude and duration. They found across a wide range of electrode configurations that this "equal-charge, equal-loudness" assumption is not valid because loudness grows more steeply from an increase in amplitude than from the same increase in duration. Their study suggests that loudness in the Nucleus device may grow unevenly as a function of the clinical units. At present, it is not clear how this uneven growth of loudness affects amplitude compression and phoneme recognition in the present study.

Another factor may be related to the difference in encoding loudness of steady-state and dynamic stimuli. At present, a steady-state test stimulus is used in all cochlear implant fitting procedures (including the Nucleus device) to measure threshold and maximal comfortable loudness levels. Recent cochlear implant studies have shown that loudness of dynamic stimuli grows differently from that of steady-state stimuli (Zeng & Shannon, 1995; Zhang & Zeng, 1997). In electric stimulation of the auditory nerve, loudness is determined by a short-term average amplitude (the root-mean-square amplitude) at the threshold level and by the peak amplitude at the maximal loudness level. Because virtually all natural sounds that implant users encounter in everyday listening situations contain fluctuating temporal envelopes, the present fitting procedure using a steady-state sound as the test stimulus will either underestimate the threshold level or overestimate the maximal loudness level for these natural dynamic stimuli. Development of standard dynamic stimuli for obtaining minimum and maximum stimulation levels may improve both quality and recognition of speech sounds.

Dynamic Range Reduction

Dawson, Skok, and Clark (1997) changed electric dynamic range by pseudo-randomly unbalancing C-levels between electrodes by 20% and found a significant effect of loudness imbalance on speech recognition. A close examination of their data (Figs. 2–5 in the Dawson et al. study) revealed that sentence recognition was more affected by loudness imbalance than phoneme recognition, recognition in noise was affected more than recognition in quiet, and MPEAK users were more significantly affected than SPEAK users. We did not collect any data in word recognition and in implant listeners using the MPEAK strategy, but our results showing a more significant effect of dynamic range reduction in noise than in quiet are consistent with the Dawson et al. study.

The present results may be limited to the SPEAK processing strategy. Loizou, Tu, and Dorman (Reference Note 3) used an acoustic simulation of the SPEAK and CIS strategy in normal-hearing listeners and found that reducing dynamic range produced a significantly more adverse effect on speech recognition in the CIS strategy than in the SPEAK strategy. This differential effect between speech processing strategies may reflect the difference in the input dynamic range as well as the type of acoustic cues extracted and delivered to the implant listeners. The SPEAK strategy employs a 30 dB acoustic dynamic range, which is significantly smaller than the approximately 50 dB range employed in the CIS strategy (Wilson et al., 1991). The SPEAK strategy extracts mostly spectral peak cues, which are well-preserved even for the binary electric dynamic range (see the third panel of Fig. 3). These spectral cues, when combined with general timing cues such as phoneme duration and gap duration, can support relatively high-level speech recognition. On the other hand, the CIS strategy generally uses fewer electrodes and explicitly extracts the channel-specific temporal-envelope information, thus being likely more sensitive to amplitude and dynamic range manipulations.

Number of Electrodes

Previous studies using an acoustic simulation of cochlear implants (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995) and in Nucleus device users (Fishman et al., 1997; Wilson, 1997) showed that in single-talker phoneme and daily sentence recognition, performance increased monotonically as the number of electrodes was increased from one to four, after which performance asymptoted. One confounding factor in the Fishman et al. study was that reducing the number of electrodes in the SPEAK strategy also increased the pulse rate per electrode, resulting in more accurate representation of the temporal envelopes. It is not clear whether and how spectral and temporal resolution can be traded with each other. However, these previous data have been taken to suggest that four electrodes are sufficient to provide the same level of performance as 20 electrodes under relatively easy listening conditions.

The present study showed that the 4-electrode processor produced significantly lower recognition scores than the 10- or 20-electrode processors in both quiet and noise conditions, particularly in noise conditions where a significant interaction was observed between the number of electrodes and phoneme recognition. One difference between the present study and previous studies is that only one talker was used in previous studies, and two talkers were used in the present study. The other difference involves speech recognition in noise in the present study. Recent studies have also found similar effects of multi-talker materials and noise on speech recognition in normal-hearing and cochlear implant listeners (Brill et al., 1997; Dorman, Loizou, & Rainey, 1997; Fu, Shannon, & Wang, in press). Together these results suggest that when the listening task becomes more challenging (such as listening to multiple talkers and in noise), more than four channels are needed to achieve optimal speech recognition performance.

Phoneme Recognition in Noise

A major advancement of the SPEAK strategy over the MPEAK strategy is its improved speech recognition in noise (Dillier, Battmer, Doring, & Muller-Deile, 1995; Skinner et al., 1994). This improvement was most apparent using sentence test materials at several signal to noise ratios. Hochberg et al. (1992) measured the recognition of phonemes in consonant-vowel-consonant words as a function of signal to noise ratios for 10 normal-hearing listeners and 10 successful implant listeners using the MPEAK strategy. Hochberg et al. defined the phoneme recognition threshold as the signal to noise ratio at which the recognition score fell to 50% of its asymptotic performance. They estimated a −2.0 dB SNR recognition threshold for the average normal-hearing listener and +10.6 dB for the average implant listener, resulting in a 12.6 dB deficit for cochlear implant listeners using the MPEAK processor.

Figure 9 shows a similar analysis of vowel (left panel) and consonant (right panel) recognition for the present normal-hearing listeners and implant listeners using the SPEAK processor. The symbols are the original recognition data re-plotted from the bottom-right panel in Figures 7 and 8 and the lines represent the fitted psychometric function (Taylor & Creelman, 1967):

$$P(X) = \alpha + (p - \alpha)\frac{1}{1 + e^{-(X-M)/S}}$$

where M is the 50% point of the asymptotic performance, S is related to the slope (about one standard deviation, or the signal to noise ratio needed to increase the performance level by 25% from the threshold), p is the asymptotic performance level, and α is the chance level performance (8.33% for vowels and 6.25% for consonants).

Table 2 presents the estimated asymptotic performance and the signal to noise ratio at which the absolute 50% correct score (the horizontal dashed line in Fig. 9) was achieved. It is clear from Figure 9 and Table 2 that cochlear implant listeners reached an asymptotic performance level that was 15 to 27% points lower than the normal-hearing listener's asymptotic performance for vowels and 28 to 33% points lower for consonants; in addition, to reach the 50% absolute correct performance level, cochlear implant listeners required 3 to 8 dB higher signal to noise ratios than normal-hearing listeners for vowels and 7 to 17 dB for consonants.

CONCLUSIONS

For cochlear implant listeners using the Nucleus SPEAK device, the amount of amplitude compression did not significantly affect phoneme recognition in either quiet or noise; the dynamic range reduction degraded phoneme recognition marginally in quiet but significantly in noise. A significant interaction between phoneme recognition in noise and the number of electrodes was also observed, suggesting that more than four electrodes are needed to optimize speech recognition in noise. Despite the success of recent speech processing development, the present study shows that a large performance gap in speech recognition still remains between cochlear implant

TABLE 2. Averaged asymptotic performance (P) in percent correct and 50% correct signal to noise ratio (SNR) in dB for normal-hearing and cochlear implant listeners (from the 20-electrode processor only).

Conditions	P (vowel)	50% SNR (vowel)	P (consonant)	50% SNR (consonant)
Normal-hearing	96.9	−9.5	94.5	−8.5
Implant-Q20	81.7	−4.1	66.0	1.5
Implant-Q50	78.0	−6.4	62.9	1.0
Implant-binary	69.5	−1.2	61.8	8.6

and normal-hearing listeners, particularly in more challenging listening conditions such as in noise. Improved cochlear implant designs and fitting procedures are required to narrow this performance gap.

ACKNOWLEDGMENTS:

The authors thank Lendra Friesen for clinical help, Mark Robert for the speech test, Chao Zhang for technical support, Nobert Dillier for the software generating the electrodogram, and Dan Huynh and John Wygonski for designing the speech-spectrum-shaped filter. We also thank Monita Chatterjee, Lendra Friesen, Qian-Jie Fu, Bob Shannon, Margo Skinner, and Alena Wilson for comments on the manuscript. This study was supported by Grant DC-02267 from National Institutes of Health.

Address for correspondence: Fan-Gang Zeng, Ph.D., Department of Hearing and Speech Sciences, University of Maryland, 0100 Lefrak Hall, College Park, MD 20742.

Received May 27, 1998; accepted October 21, 1998

REFERENCES

Boothroyd, A., Erickson, F. N., & Medwetsky, L. (1994). The hearing aid input: A phonemic approach to assessing the spectral distribution of speech. *Ear and Hearing, 6*, 432–442.

Brill, S. M., Gstottner, W., Helms, J., Ilberg, C. V., Baumgartner, W., Muller, J., & Kiefer, J. (1997). Optimization of channel number and stimulation rate for the fast Continuous Interleaved Sampling strategy in the COMBI 40+. *American Journal of Otology*, 18, S104–S106.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Nasser Kotby, M., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G., Westerman, S., & Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *Journal of Acoustical Society of America, 96*, 2108–2120.

Borg, E., & Zakrisson, J.-E. (1973). Stapedius reflex and speech features. *Journal of Acoustical Society of America, 54*, 525–527.

Cochlear Corporation (1995). *Technical reference manual*. Englewood, Colorado.

Dawson, P. W., Skok, M., & Clark, G. M. (1997). The effect of loudness imbalance between electrodes in cochlear implant users. *Ear and Hearing, 18*, 156–165.

Dillier, N., Battmer, R. D., Doring, W. H., & Muller-Deile, J. (1995). Multi-centric field evaluation of a new speech coding strategy for cochlear implants. *Audiology, 34*, 145–159.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America, 102*, 2403–2411.

Eddington, D. K. (1980). Speech discrimination in deaf subjects with cochlear implants. *Journal of Acoustical Society of America, 68*, 885–891.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, D. E., & Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *Annals of Otology, Rhinology, and Laryngology, 87*(Suppl. 53), 5–39.

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

Fletcher, H. (1953). *Speech and hearing in communication*. New York: D. Van Nostrand Company, Inc.

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

Fu Q.-J., Shannon, R. V., & Wang, X. (in press). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *Journal of the Acoustical Society of America.*

Hickson, L., & Byrne, D. (1997). Consonant perception in quiet: Effect of increasing the consonant-vowel ratio with compression amplification. *Journal of the American Academy of Audiology, 8*, 322–332.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of 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.

Hudspeth, A. J. (1997). How hearing happens. *Neuron*, 19, 947–950.

House, W. F. (1978). The clinical value of single electrode system in auditory prostheses. *Otolaryngology Clinicians of North America, 11*, 201–208.

Lippmann, R. P., Braida, L. D., & Durlach, N. I. (1981). Study of multi-channel amplitude compression and linear compression for persons with sensorineural hearing loss. *Journal of the Acoustical Society of America, 69*, 524–534.

McDermott, H. J., McKay, C. M., & Vandali, A. E. (1992). A new portable sound processor for the University of Melbourne/ Nucleus Limited Multichannel cochlear implant. *Journal of Acoustical Society of America, 91*, 3367–3371.

Merzenich, M. M., Rebscher, S. J., Loeb, G. E., Byers, C. L., & Schindler, R. A. (1984). The UCSF cochlear implant project. State of development. *Advances in Audiology, 2*, 119–144.

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. Supplement, 166*, 303–306.

Nelson, D. A., Schmitz, J. L., Donaldson, G. S., Viemeister, N. F., & Javel, E. (1996). Intensity discrimination as a function of stimulus level with electric stimulation. *Journal of Acoustical Society of America, 100*, 2393–2414.

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

Shannon, R. V., Zeng, F.-G., & Wygonski, J. (1992). Speech recognition using only temporal cues. In M. E. H. Schouten (Ed.), *The auditory processing of speech: From sounds to words* (pp. 263–274). Berlin: Mouton de Gruyter.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P. M., Staller, S. J., Shipp, D. B., Shallop, J. , 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. *The American Journal of Otology, 15*(Suppl. 2), 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., & 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.

Souza, P. E., & Turner, C. W. (1996). Effect of single-channel compression on temporal speech information. *Journal of Speech and Hearing Research*, *39*, 901–911.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of Acoustical Society of America*, *41*, 782–787.

Turner, C. W., Souza, P. E., & Forget, L. N. (1995). Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *Journal of Acoustical Society of America*, 97, 2568–2576.

Van Tasell, D. J., & Trine, T. D. (1996). Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. *Journal of Speech and Hearing Research*, *39*, 912–922.

Viemeister, N. F. (1988). Intensity coding and the dynamic range problem. *Hearing Research*, *34*, 267–274.

Wilson, B. S. (1997). The future of cochlear implants. *British Journal of Audiology*, *31*, 205–225.

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., & Zhang, C. Y. (1998). Encoding loudness by electric stimulation of the auditory nerve. *NeuroReport*, *9*, 1845–1848.

Zeng, F.-G., & Shannon, R. V. (1992). Loudness balance between acoustically and electrically stimulated ears. *Hearing Research*, *60*, 231–235.

Zeng, F.-G., & Shannon, R. V. (1994). Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. *Science, 264*, 564–566.

Zeng, F.-G., & Shannon, R. V. (1995). Loudness of simple and complex stimuli in electric hearing. *Annals of Otology, Rhinology and Laryngology*, *104*(Suppl. 166), 235–238.

Zeng, F.-G., Shannon, R. V., & Hellman, W. S. (1998). Physiological processes underlying psychophysical laws. In A. R. Palmer, A. Rees, A. Q. Summerfield, & R. Meddis (Eds.), *Psychophysical and physiological advances in hearing* (pp. 473–481). London: Whurr Publishers Ltd.

Zeng, F.-G., & Turner, C. W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired listeners. *Journal of Speech and Hearing Research*, *33*, 440–449.

Zhang, C. Y., & Zeng, F.-G. (1997). Loudness of dynamic stimuli in acoustic and electric hearing. *Journal of Acoustical Society of America*, *102*, 2925–2934.

REFERENCE NOTES

1 Boex, C. S., Eddington, D. K., Noel, V. A., Rabinowitz, W. M., Tierney, J., & Whearty, M. E. (1997). Restoration of normal loudness growth for CIS sound coding strategies. Paper presented at the Conference on Implantable Auditory Prostheses, Pacific Grove, California.

2 Cosendai, G., & Pelizzone, M. (1997). Acoustic dynamic range of compressive mapping and speech recognition with cochlear implants. Paper presented at the Conference on Implantable Auditory Prostheses, Pacific Grove, California.

3 Loizou, P., Tu, Z., & Dorman, M. (1998). The intelligibility of speech composed of varying number of sinusoids. Paper presented at the IEEE International Conference on Acoustics, Signal and Speech Processing, Seattle, Washington.

4 Svirsky, M. A., & Meyer, T. A. (1998). A mathematical model of consonant perception by cochlear implant users with the SPEAK strategy. Paper presented at the 16th ICA and 135th ASA meeting, Seattle, Washington.