RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 25 (2001-2002)
Indiana University

Perception of "Elliptical Speech" Following Cochlear Implantation: Use of Broad Phonetic Categories in Speech Perception[1]

Steven Fountain[2] and David B. Pisoni[2]

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This work was supported by NIH T32 Training Grant DC00012 to Indiana University. We thank Luis Hernandez for computer programming and support in this experiment, Rebecca Herman who voiced half and created all the stimuli used and Michael Vitevitch who voiced half of the stimuli used in this experiment. We also thank the ENT Hearing and Speech Clinic for providing the cochlear implant patient list and the setting where the experiment took place.
[2] DeVault Otologic Research Laboratory, Department of Otolaryngology, Indiana University School of Medicine, Indianapolis, IN.

Perception of "Elliptical Speech" Following Cochlear Implantation: Use of Broad Phonetic Categories in Speech Perception

Abstract. The present study investigated the perception of "elliptical speech" (Miller & Nicely, 1955) in 10 hearing-impaired adults with cochlear implants. A group of 15 normal-hearing adults were also tested for comparison. Sentence discrimination and repetition tasks were employed using sets of meaningful and anomalous English sentences. Two different versions of each set of sentences were constructed. One version contained sentences with intact place of articulation cues; the other version transformed the sentences into "elliptical speech" using a procedure in which the consonants were all replaced with other consonants that had the same voicing and manner of articulation features but always had an alveolar place of articulation. The hearing-impaired listeners who use cochlear implants completed a same-different sentence discrimination task and a repetition task under quiet listening conditions. The normal-hearing listeners completed both tasks under noise-masking conditions. In the same-different discrimination task, normal-hearing listeners perceived sentences with intact place of articulation cues and its elliptical version as the "same." The cochlear implant users showed a similar response pattern. These findings support Miller and Nicely's claim that under signal degradation conditions, ellipsis can no longer be detected reliably. In the repetition task, however, normal-hearing subjects showed somewhat better repetition performance for sentences with intact place of articulation than for elliptical speech sentences. This was unexpected given the earlier findings from the sentence discrimination task. The cochlear implant users also showed slightly better repetition performance for sentences with intact place of articulation cues than for elliptical speech. Taken together, both sets of findings on the perception of elliptical speech provide support for the hypothesis that hearing-impaired patients with cochlear implants perceive speech and recognize spoken words using broad perceptual equivalence classes. Even without highly detailed acoustic-phonetic information about place of articulation in speech, many patients with cochlear implants are able to reliably recognize words in sentences and repeat them immediately on-the-fly after only one presentation.

Introduction

What does speech sound like to a deaf patient with a cochlear implant? This is an important unexplored question in speech perception and spoken word recognition that has both theoretical and clinical implications. Do patients hear speech as a sequence of meaningful spoken words arrayed in time or do they perceive and recognize the degraded and impoverished nature of the speech signals that are transmitted by their cochlear implant? Clinical experience suggests that listeners who use cochlear implants often do not perform well on open-set tests of word recognition that require them to identify spoken words by accessing all possible words in their entire mental lexicon. Many of the confusions shown by deaf listeners who use cochlear implant in these word recognition tests are due to the problems in perceiving the place of articulation of consonants. However, despite this apparent difficulty in correctly perceiving place of articulation, many cochlear implant users are able to do very well in face-to-face conversation and other tasks that require comprehension of the intended message.

In their pioneering work on speech perception, Miller and Nicely (1955) found that place of articulation was frequently confused under masking and low-pass filtering. Based on their finding that some consonants tended to be confused with each other, Miller and Nicely grouped consonants into perceptual equivalence classes that they assumed were functionally the same under these degraded listening conditions. For example, if [p t k] are mistaken for one another frequently under degraded listening conditions, these segments would be placed in a single equivalence class. Miller and Nicely further suggested that a unique form of phonetically degraded speech could be created in which a single member of each equivalence class replaced every other individual member of its equivalence class. They called this kind of speech "elliptical speech" because of the ellipsis, or omission, of place of articulation information. For example, if speech is produced in which every [p t k] was simply replaced by [t], the resulting speech sounds very strange when presented in the clear. However, when this speech is presented under degraded listening conditions where the confusions were originally produced, Miller and Nicely predicted that elliptical speech should actually be undetectable because the members of the equivalence classes are grouped precisely under the same signal degradation conditions. Miller and Nicely reported some informal support for their predictions, although they never carried out a formal experiment to support their hypothesis (Miller, 1956; Miller & Nicely, 1955).

Recently, Quillet, Wright and Pisoni (1998) reported the results of a same-different discrimination experiment that was designed to assess the predictions made by Miller and Nicely that elliptical speech would be undetectable under conditions of signal degradation. Using synthesized speech, young normal-hearing listeners were presented with pairs of sentences that were either lexically the same or lexically different. In one condition, both sentences in a lexically identical pair were intact and were presented without any ellipsis. In a second condition, both sentences in a pair were transformed into elliptical speech. Finally, in a third condition, one sentence in a lexically identical pair was presented intact while the other was an elliptical version. Quillet et al. found that when the stimuli were presented in the clear with no signal degradation, the listeners identified most of sentences in this third condition as "different." However, under signal degradation conditions, the majority of the listeners identified a large number of the pairs in this condition as "same" indicating that the ellipsis of place of articulation could not be detected under these conditions.

In addition to replicating Miller and Nicely's informal experiment on elliptical speech and providing support for their earlier predictions, Quillet et al. (1998) also noted several interesting parallels in speech perception between normal-hearing listeners under conditions of signal degradation and deaf listeners who use cochlear implants. Just as normal-hearing listeners show systematic confusions among different places of articulation under conditions of signal degradation, patients with cochlear implants also show similar confusions among places of articulation. Quillet et al. suggested that it might be possible to use "elliptical" speech as a research tool to investigate the perception of speech by patients with cochlear implants and to try to understand how they do so well with their implant even with highly impoverished input signals. If we find that cochlear implant users perceive "elliptical speech" as the same as intact speech, this result would provide support for the hypothesis that cochlear implant users employ broad perceptual equivalence classes to perceive place of articulation in speech perception.

The present investigation is an extension of a recent case study conducted by Herman and Pisoni (2000). They tested an exceptionally good cochlear implant patient ("Mr. S") under conditions of elliptical speech and 20 normal-hearing listeners using degraded speech signals. The design of the present experiment was similar to the preliminary study carried out by Herman and Pisoni. In both studies, a same-different discrimination task was employed to determine if elliptical speech would be undetectable to listeners who use cochlear implants. The two earlier investigations differed in terms of how the stimuli were created. Herman and Pisoni created their stimuli using natural speech whereas Quillet et al. used synthesized speech.

In the present study, pairs of English sentences were presented to adult cochlear implant users who were asked to determine whether the two sentences were the "same" or "different." As in Quillet et al., the crucial test case was the third condition in which two sentences were lexically identical, but one was intact while the other was transformed into elliptical speech. In this condition, under the hypothesis that cochlear implant users cannot detect place of articulation information, we predicted that our patients would label the two sentences as the "same." This pattern of results would suggest that consonants with the same manner and voicing features but different places of articulation form an equivalence class and that cochlear implant users recognize words in context using broad phonetic categories. This pattern of results would also provide support for the hypothesis that cochlear implant users perceive speech as a sequence of highly familiar words and do not normally detect fine acoustic-phonetic differences between spoken words in meaningful sentences.

Up to this point, we have focused on what speech might sound like to users of cochlear implants, and thus what obstacles might have to be overcome to recognize words successfully. A second question of interest to us concerns why some users of cochlear implants manage to perceive speech so well in face-to-face conversations despite the degraded auditory signals they receive through their implants. One explanation for their good performance in the real world is the observation that powerful structural constraints influence the sound patterns found in all languages (Shipman & Zue, 1982; Zue & Huttenlocher, 1983). For example, Zue and Huttenlocher (1983, p. 122) observed that the sound patterns of words in spoken languages are highly constrained not only by the inventory of individual speech sounds in a particular language but also by the allowable combinations of those sound units" (i.e., the phonotactic constraints). Shipman and Zue (1982) reported that an analysis of English that distinguishes only between consonants and vowels can prune a 20,000-word lexicon down to about 200 CV combinations.

Because strong structural constraints on sound patterns exist in language, a broad phonetic classification can serve to define a "cohort" or the set of possible candidate words having the same pattern. As Shipman and Zue showed in their computational research, these candidate sets may actually be quite small. They found that the average size for these equivalence classes for the 20,000-word lexicon was approximately 2 and the maximum size was approximately 2,000 (Zue & Huttenlocher, p. 122). Thus, even if a listener does not accurately perceive the exact place of articulation, he can still identify words successfully using broad equivalence classes if he can recognize at least the sequence of consonants and vowels in the pattern. Broad phonetic coding of the input signal may be sufficient to permit higher-level lexical processes to take over and support word recognition and spoken language comprehension despite degraded input signals.

While research has shown that phonotactic constraints exist in language, listeners may not necessarily use them in real-time speech perception. Quillet et al. (1998) investigated whether coarse coding of the speech signal could provide a rich and sufficient set of cues to allow normal-hearing listeners to understand meaningful sentences. In a transcription task using synthesized speech, young normal-hearing listeners were asked to transcribe key words from each sentence. The sentences had either intact place of articulation cues or were produced using elliptical speech. The sentences were presented in the clear or in white noise at 0 dB SNR, -5 dB SNR, and -10 dB SNR. Quillet et al. predicted that while speech with intact place of articulation should show decreased intelligibility under degraded conditions, elliptical speech should actually show the reverse pattern, that is, increased intelligibility as signal degradation increased. Quillet et al. found that speech with intact place of articulation cues did show decreases in transcription accuracy under degraded conditions whereas the elliptical speech showed improvements in transcription accuracy from the 0 dB SNR level to the -5 dB SNR level before dropping off at the -10 dB SNR level. Quillet et al. interpreted these findings as support for the proposal that normal-hearing listeners are able to make use of broad phonetic categories to identify spoken words in sentences under conditions of signal degradation.

In their case study of "Mr. S," Herman and Pisoni (2000) explored whether one excellent cochlear implant user utilized coarse coding and broad phonetic categories in perceiving sentences. They used a transcription task that was similar to the task used by Quillet et al. However, Herman and Pisoni used natural speech rather than synthetic speech. They predicted that normal-hearing listeners and their cochlear implant patient would transcribe elliptical speech at the same level of accuracy as intact speech.

Herman and Pisoni found that the normal-hearing listeners and their cochlear implant patient transcribed intact speech more accurately in than elliptical speech, a result that differed from Quillet et al.'s earlier findings. The pattern of results indicates that some conflicting phonetic cues to place of articulation were actually perceived in the stimuli. These cues created confusions for the listeners in recognizing the elliptical speech samples. To explain the discrepancy between the two studies, Herman and Pisoni suggested that the redundant speech cues present in the natural speech stimuli might have survived the signal degradation more than Quillet et al.'s synthetic speech, which was much less redundant than the natural speech samples. Also, Herman and Pisoni provided their listeners with the option of repeating the sentence up to five times before they had to record their responses. This may also have improved their performance under the degraded conditions. In Quillet et al., listeners only heard each test sentence once.

The main difference between the earlier study by Herman and Pisoni (2000) and the present investigation is the format of the discrimination task. In Herman and Pisoni, a transcription task was employed in which the participants simply wrote down "keywords" of a sentence presented to them over a loudspeaker. For each sentence, a response frame with blanks substituted for the keywords was provided for them to transcribe their answers. Additionally, the participants had the option of listening to the sentence up to five times in order to fill in each blank. The methodology was changed for this study to a repetition task. Listeners were asked to repeat out loud as much of the sentence as possible after listening to the sentence only once. This was thought to be a better way to measure of the immediate perception of the sentences. Unlike the earlier transcription task, the participants did not have access to the redundancy of the phonetic cues that may have been utilized by listeners who could hear the sentence repeatedly up to five times.

Although these differences may explain part of the discrepancy in the results between Herman and Pisoni and Quillet et al., natural speech was still used to create conditions that better resembled the real world. Because of the difficulty in carrying out an immediate repetition task, only cochlear implant users who scored above a minimum satisfactory level in a standard open-set word recognition test were selected for this study. Again, half of the sentences were intact and half were transformed into elliptical speech. If the coarse-coding hypothesis of spoken word recognition is correct, cochlear implant users should show the same performance on sentences with intact place of articulation cues as on sentences transformed into elliptical speech. This pattern would indicate that coarse coding was sufficient for successful word recognition to be carried out in spoken sentences.

Method

Participants

Ten severe-to-profoundly deaf patients who use cochlear implants were recruited based on their age, length of time using their implant, and most recent CNC list score from the patient charts at the Department of Otolaryngology, Indiana University School of Medicine. All of the patients were postlingually deafened and had acquired language normally before the onset of their hearing loss. Each participant with a cochlear implant was between the ages of 24 and 71 and had used their implant for at least one year. A standard clinical measurement of speech perception performance, their CNC score, represents the percentage of words that a hearing-impaired person correctly recognizes immediately following presentation of each word in an open-set format. A CNC score of 30% was the criterion used for participation in this study. Table 1 shows a summary of the demographic details of the patients who used cochlear implants.

In addition to the 10 patients, 15 normal-hearing listeners between the ages of 21 to 40 were also recruited using an email distribution list at the IUPUI campus. Both students and staff of the university responded and received a payment of 10 dollars for their time participating in the experiment. These listeners received the same conditions as the cochlear implant patients except that half of their trials were run under degraded listening conditions. None of the listeners reported any hearing or speech problems at the time of testing. All participants were native speakers of American English.

Cochlear implant user	Years with implant	CNC word score
1	2	62
2	6	50
3	2	86
4	5	30
5	3	46
6	1	77
7	3	50
8	5	74
9	1	38
10	10	64

Table 1. Age, number of years with implant, and CNC score of each cochlear implant user.

Stimulus Materials

The stimulus materials used in this study consisted of 96 Harvard sentences (IEEE, 1969). Each sentence contained five key words with declarative or imperative sentence structure, taken from lists 1-10 of Egan (1948). In addition to the *normal* Harvard sentences, a set of 96 *anomalous* sentences were created by substituting random words of the same lexical category (noun, verb, etc.) into the Harvard sentences of lists 11-20 (Egan, 1948). The anomalous sentences were developed to block normal top-down semantic processing. The new words were selected from lists 21-70 of the Harvard sentences (Egan, 1948). New sets of "elliptical" sentences were generated using these two types of sentences through a process of featural substitution that was similar to the original procedures developed by Miller and Nicely (1955). The stops, fricatives, and nasal consonants in each of the five key words were replaced with a new consonant that preserved the same manner and voicing features of the original consonant but changed the place feature to an alveolar place of articulation. For example, "See the plane in the blue sky" would be changed to "See the tlane in the dlue sky." Liquids /r l/ and glides /y w/ were excluded from the substitution process. Intact sentences and their elliptical versions are listed in the appendix along with the equivalence classes of consonants. This method of replacing consonants with alveolar consonants follows Miller and Nicely's original method of creating elliptical speech and differs from the procedures used earlier by Quillet et al. (1998).

A male and a female talker were used to generate the stimulus materials. Each talker read half of the sentences aloud. Before the recording session, both talkers practiced saying the test sentences several times. The speaker attempted to use the same intonation pattern in both the intact and elliptical versions of an utterance. Sentences were recorded using a head-mounted Shure Model SM98A microphone and a Sony TCD-D8 DAT recorder. The recordings were segmented into individual utterances, converted to a single channel, and downsampled to 22,050 Hz using CoolEdit™.

For the signal degradation, a masking noise was created using Gaussian noise applied to each sentence to create another set of stimuli. Noise was added at a -5 dB signal-to-noise ratio. Each noise-masked file was then saved as a separate digital speech file for use during presentation of the stimuli to the listeners.

Procedures

Sentence Discrimination Task. The 10 cochlear implant patients heard the test stimuli over an AV570 Advent loudspeaker. They were given four practice trials in which they could adjust the volume of the speaker to a comfortable listening level. A Visual Basic program running on a PC recorded subject responses and controlled the experiment. The experiment was self-paced to each subject. Each pair of sentences was presented only once. There was a one-second interval between the two sentences in each pair. Responses were entered using the computer mouse to click on a dialog box labeled "same" or "different" on the computer monitor. The participants were given explicit instructions to use the label "same" only for pairs of sentences that sounded exactly word-for-word and sound-for-sound identical. Each cochlear implant patient was presented with 64 sentences in two blocks of 32 trials each. They heard a block of normal sentences followed by a block of anomalous sentences. Half of the sentences in each block were spoken by a male speaker and half by a female speaker. Likewise, half of the sentences in each block were composed of intact speech cues and half consisted of elliptical speech sentences. A Visual Basic randomizer program on the PC was employed to randomize the order of the stimuli in each block.

Fifteen normal-hearing listeners followed the same procedures except that the volume of the speaker was set at a comfortable level of 80 dB. Also, the normal-hearing listeners did not receive the sentences in separate blocks but heard them in random order. Since normal-hearing listeners received half of the sentences under degraded listening conditions, 128 pairs of sentences, twice the number that cochlear implant users received, were presented to them. This procedure insured that the same number of sentences would be presented in the clear to both the cochlear implant patients and the normal-hearing listeners. Sentences presented in the clear and under degraded conditions were presented to normal-hearing listeners in a random order.

	Lexically Different	Lexically Identical
Both intact speech	IaIz	IaIa
Both elliptical speech	EaEz	EaEa
One intact, one elliptical	IaEz EaIz	IaEa EaIa

Table 2. The different pairs of sentences used in the sentence discrimination task.

All participants received eight pairs of test sentences, as shown in Table 2. Pairs of sentences that were lexically identical are marked with two subscript "a's." Pairs of sentences that were lexically different are marked with a subscript "a" and a subscript "z." The sentences with intact place of articulation cues are referred to as "I" while the elliptical sentences are referred to as "E."

Sentence Repetition Task. The second part of the experiment employed an immediate repetition task. Subjects heard a sentence one time and were asked to immediately repeat back as much of the sentence as possible. Different sets of sentences were used in this task to avoid repetition and priming effects from earlier sentences used in the discrimination task. A Visual Basic program running on a PC controlled the experiment. This experiment was also self-paced. After repeating a sentence, the participant used a mouse to press a "next sentence" dialog box on the computer monitor to advance to the next trial. A microphone connected to a tape recorder recorded the subjects' vocal responses on each trial.

Ninety-six sentences were presented. Half of the sentences were normal sentences and half were anomalous sentences. Likewise, half of the sentences in each set contained intact speech cues and half contained elliptical speech cues; half were spoken by the female talker and half by the male talker. Normal-hearing participants heard half of their sentences in noise and half in the quiet. The patients with cochlear implants heard all of their sentences in the quiet.

Responses were scored using a very strict criterion. Responses were scored as correct if and only if they exactly matched the intended word. In the elliptical speech conditions, a word was scored as correct if it was produced as the original English word and incorrect if it was produced with any of the degraded speech cues that were actually heard (i.e., the elliptical version). For example, if the target word was "blue" and the elliptical version that was heard in the sentence was "dlue," a response of "dlue" would be scored as "incorrect." If the subject said "blue," it would be scored as "correct." A score for each sentence was computed based on the correct repetition of the three keywords in each sentence. A score of 0,1,2, or 3 was assigned if no keyword, one keyword, two keywords, or three keywords respectively were "correctly" repeated.

Results

Sentence Discrimination

Normal-hearing Listeners. The average number of "same" responses from the 15 normal-hearing listeners for both normal and anomalous sentences is shown in the two panels in Figure 1. Sentences presented in the clear are shown by light bars; sentences presented under degraded conditions are shown by dark bars. The different sentence pairings (i.e., IaIz, EaEz) are presented on the abscissa.

Normal-hearing listeners performed as expected on both normal and anomalous sentences. For the first four pairs of sentences in which the two sentences were lexically different, almost all of the pairs were labeled "different." For the next two sentence pair types, in which the two sentences were exactly identical, nearly all pairs were labeled "same." In the critical conditions, the last two sentence pair types, in which the two sentences were lexically the same although one sentence contained intact speech cues and the other sentence contained elliptical speech cues, normal-hearing listeners perceived most of them as "different" when they were presented in the clear. However, when the sentences were presented under degraded conditions, normal-hearing listeners could not discriminate between intact speech and elliptical speech versions of the same sentences and responded "same" on almost all of the trials.

The responses from of the normal-hearing listeners were subjected to a 2x2x4 analysis of variance (ANOVA). The first factor was "meaningfulness." The two levels were normal vs. anomalous sentences. The second factor was signal degradation. The two levels were sentences heard in clear and sentences heard under noise masking. The third factor was the pair type. Four levels for pair type are shown in the different cells in Table 2 and are plotted on the abscissa in Figure 1: IaIz and EaEz (two different sentences, either both have intact speech cues or both have elliptical speech cues), IaEz with EaIz (two different sentences, one has intact speech cues and one has elliptical speech cues), IaIa with EaEa (the same sentence twice, both have intact speech or both have elliptical speech cues), and the crucial test cases of IaEa with EaIa (the same sentence twice, one with intact speech cues and one with elliptical speech cues).

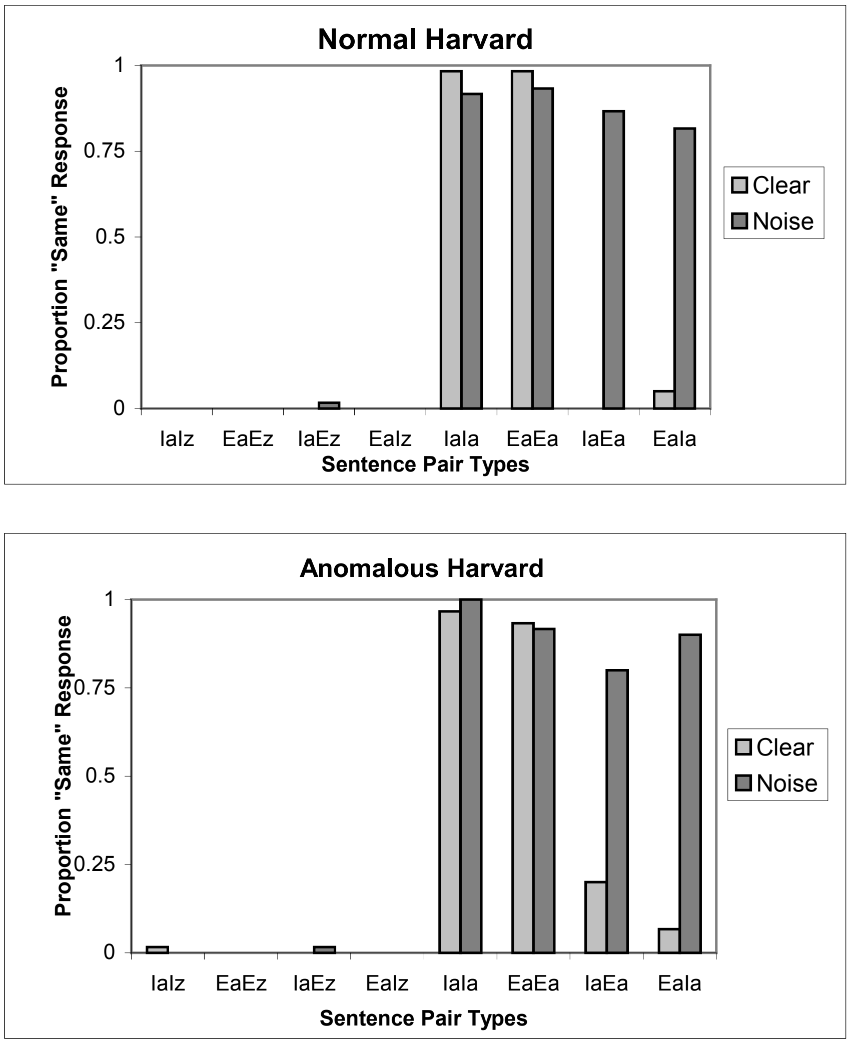


Figure 1. Results from the sentence discrimination task for normal-hearing listeners.

A significant difference in discrimination performance was obtained between sentences heard in the clear and sentences heard under degraded conditions (F (1,14) = 159.8, p <.001). The normal sentences and anomalous sentences did not differ significantly. The analysis also showed a significant two-way interaction between the signal degradation and pair type. Broken down by level using post-hoc tests for simple effects, signal degradation within the first level (IaIz and EaEz) and second level (IaEz and EaIz) showed no differences. When normal-hearing listeners heard two lexically different sentences, they had no difficulty with these pairs of sentences and were able to correctly discriminate these differences and respond appropriately regardless of whether the sentences contained both intact speech cues or both elliptical speech cues, or whether one sentence was intact and one was elliptical. Also, we found no difference in performance between sentences presented in the clear and sentences masked by noise. Within the third level (IaIa and EaEa) where identical sentences were presented twice, no difference was observed. Listeners correctly labeled the two sentences as the "same" most of the time. Again, we observed no difference in performance between the sentences presented in the clear and sentences masked by noise. These findings are consistent with our earlier predictions.

We did observe a significant difference for the "same" responses under signal degradation for the fourth set of sentences (IaEa and EaIa ($F(1,14) = 160.7$, $p < .001$)). For these critical pairs of sentences, two lexically identical sentences were presented. One sentence was intact and the other contained elliptical speech cues. In the clear, listeners correctly labeled almost all of these sentence pairs as "different," but under degraded listening conditions, almost all of the sentence pairs were labeled as "same." This is exactly the pattern of responding we expected to find if the normal-hearing listeners used broad equivalence classes to recognize spoken words in these sentences. And, this is precisely the pattern that would be predicted by Miller and Nicely based on the consonant confusions generated in noise.

Cochlear Implant Patients

A summary of the results for the 10 cochlear implant patients is shown in Figure 2. The results from the patients look quite similar to the data obtained from the normal-hearing listeners under degraded listening conditions. Lexically different sentence pairs were labeled "different" almost all of the time while lexically identical sentences were labeled "same" most of the time. The crucial test conditions were the last two sentences in which two lexically identical sentences were presented with one containing intact speech cues and the other containing elliptical speech cues. The results demonstrate that cochlear implant users for the most part could not discriminate differences between intact speech cues and elliptical speech cues. They perceived both types of elliptical sentences as the same more than 75% of the time.

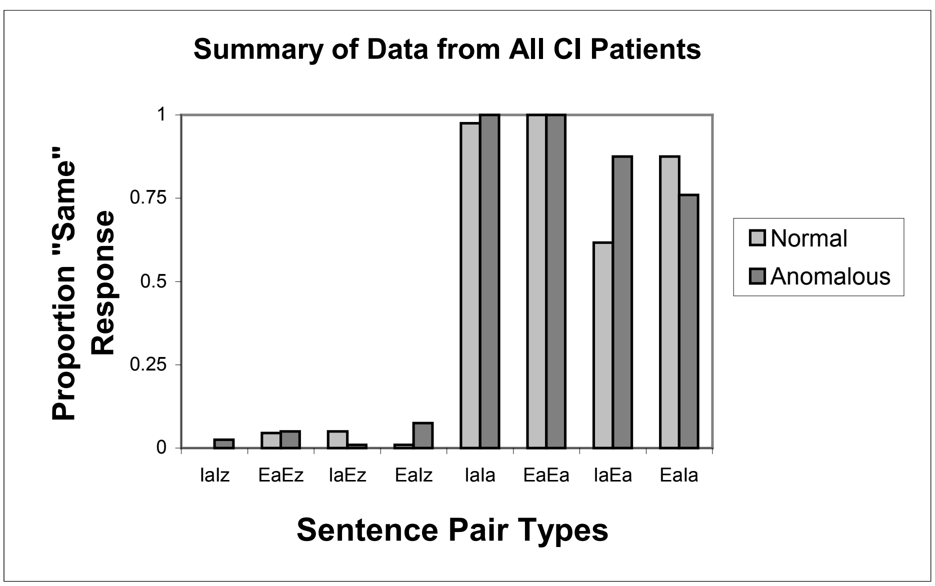


Figure 2. Mean results for the 10 cochlear implant patients.

The data were analyzed using a 2x4 ANOVA with the factors "meaningfulness" and "pair type" as in the normal-hearing listener's data analysis but without the additional factor of "degradation." No differences in discrimination were found between the normal and anomalous sentences.

To examine individual patient's data, three groups of listeners were created based on the earlier CNC word recognition scores. CNC scores from 30 to 40 were designated "fair" users, CNC scores from 41 to 62 were designated "good" users, and CNC scores from 63 to 86 were designated "excellent" users. Figure 3 shows individual data for the two fair users. Figure 4 shows individual data for the four good users. Figure 5 shows individual data for the four excellent users. Each patient's CNC score is shown next to his/her subject number in the legends of each figure. In each figure, normal sentences are shown in the top panel while anomalous sentences are shown in the bottom panel.

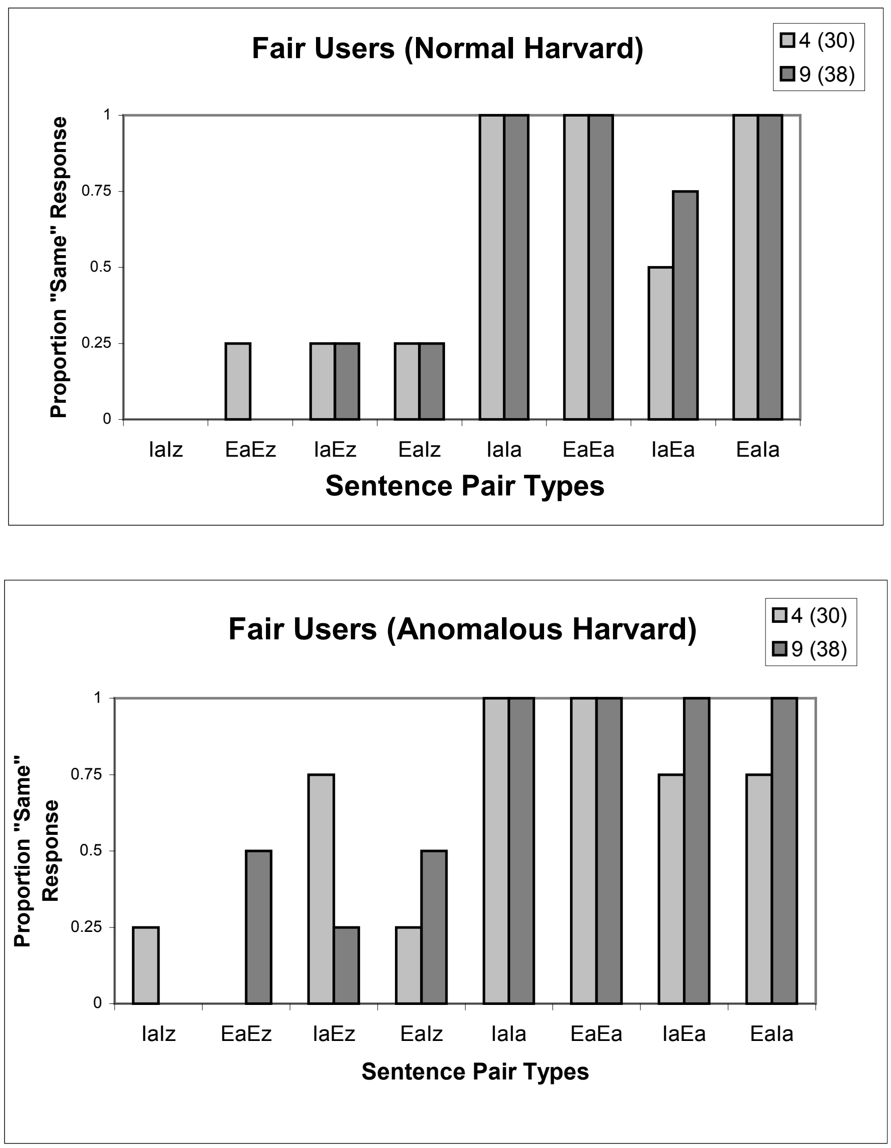
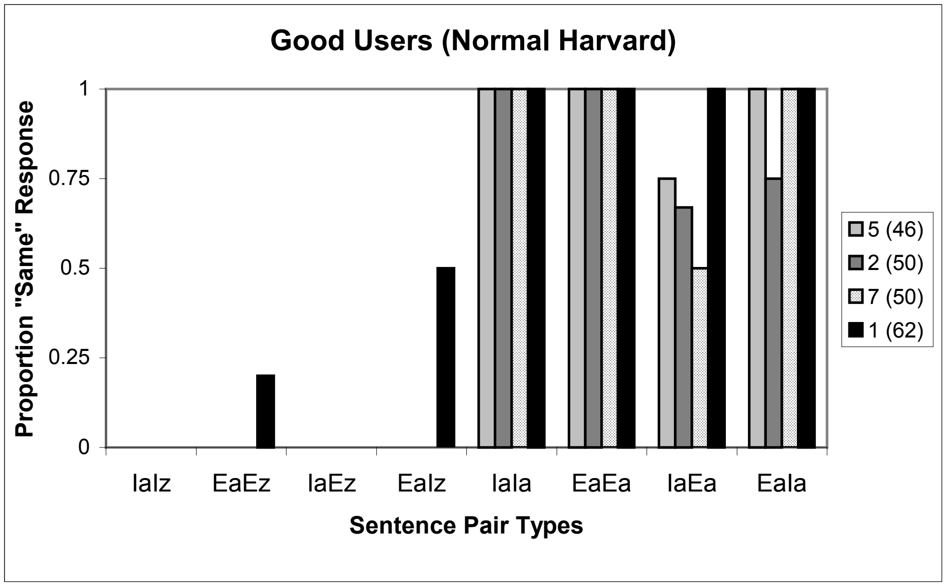


Figure 3. Individual results for the fair CI users.



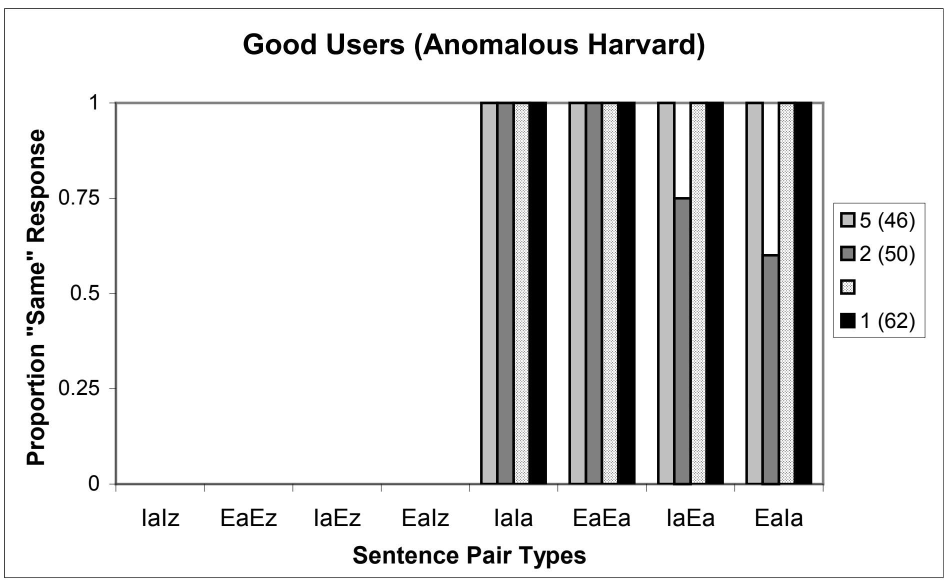
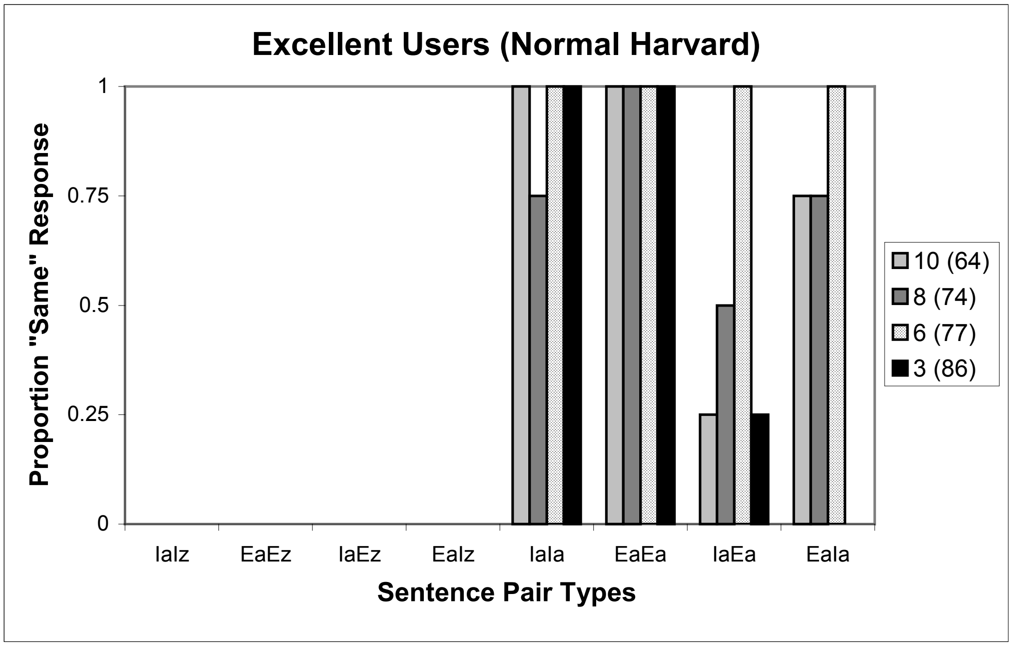


Figure 4. Individual results for the good CI users.

While the three groups of implant patients displayed response patterns that were similar to normal-hearing listeners under degraded listening conditions, the two extreme groups, the “fair” and “excellent” users showed somewhat different patterns of discrimination. The “fair users” had a great deal of difficulty distinguishing even between sentences that were lexically different. This is not surprising because they had low CNC scores to begin with. The “good users” rarely made that kind of error and the “excellent users” were always able to discriminate differences between lexically different sentences. On the other hand, the “excellent” users were able to discriminate differences between intact and elliptical speech and they performed better than the “good” users. The best user, Patient #3, with a CNC score of 86 was consistently able to discriminate between intact and elliptical speech cues in these sentences.



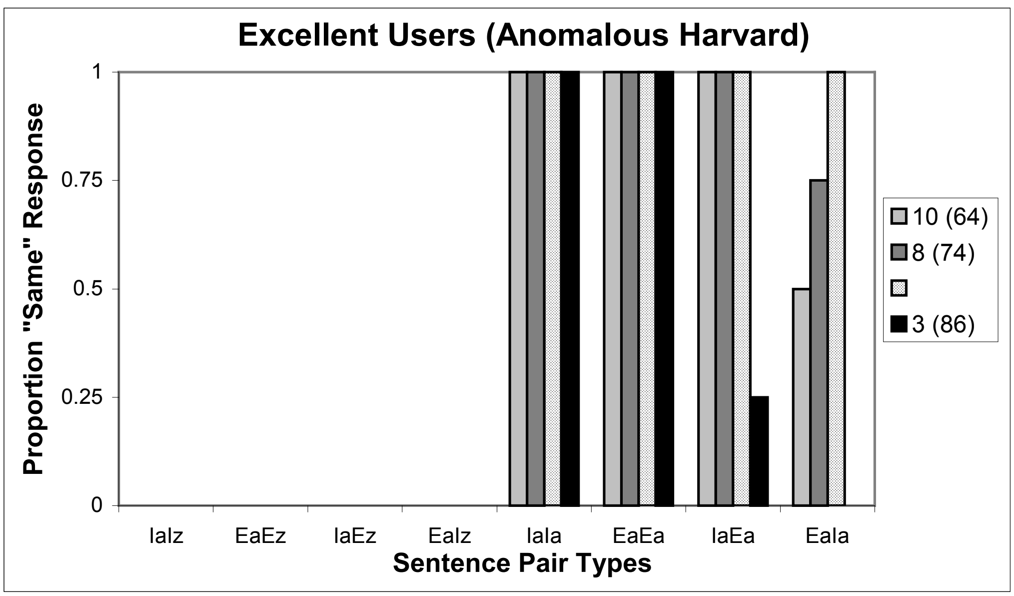


Figure 5. Individual results for the excellent CI users.

Sentence Repetition Task

Normal-hearing Listeners. The average correct proportion repetition performance for the 15 normal-hearing listeners for both normal and anomalous sentences is shown in Figure 6. Sentence Type is shown along the abscissa. The normal sentences are shown on the left; the anomalous sentences are shown on the right. Proportion correct indicates the proportion of intact word responses. Sentences presented in the clear are represented by the light bars and sentences presented under noise masking are shown by the dark bars.

The pattern of results is consistent with our earlier expectations. In the clear, normal-hearing listeners repeated words much better from intact sentences than elliptical sentences. This is due to our method of scoring since normal-hearing listeners repeated the words that were presented to them. Thus, in the clear, a sentence containing elliptical speech cues would be repeated with the incorrect speech cues present. However, for sentences in noise, normal-hearing listeners scored about the same for sentences containing intact speech cues and elliptical speech cues.

A 2x2x2 ANOVA was used to assess differences in the scores. The three factors were (a) speech cues (intact vs. elliptical), (b) signal degradation (clear vs. noise), and (c) meaningfulness (normal vs. anomalous sentences). Significant main effects were found for each factor: (a) F (1,14) = 1889.6, p<.001, (b) F (1,14) = 1071.3, p<.001, (c) F (1,14) = 15.6, p = .001, respectively.

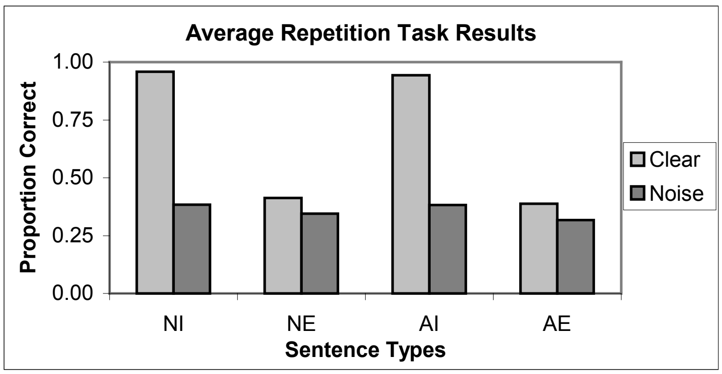


Figure 6. Average results from the repetition task for normal-hearing listeners. NI =normal Harvard intact sentence, NE =normal Harvard elliptical sentence, AI =anomalous Harvard intact sentence, AE =anomalous Harvard elliptical sentence.

For normal sentences, a separate 2x2 ANOVA revealed a significant main effect of speech cues (F (1,14) = 1686.6, p<.001). A significant main effect was also observed for signal degradation (F (1,14) = 473.9, p<.001). A post-hoc test for simple effects showed a significant difference between intact and elliptical speech cues for the normal sentences heard in noise (F (1,14) = 8.38, p< 0.5). For the anomalous sentences, a 2x2 ANOVA showed a significant difference for both speech cues (F (1,14,) = 935.3, p<.001) and signal degradation (F (1,14) = 841.1, p<.001). A post-hoc test revealed a significant difference (F (1,14) = 44.6, p = .001) between intact and elliptical speech of anomalous sentences heard in noise. Although normal-hearing listeners under degraded conditions appeared to score about the same for sentences containing intact speech cues and sentences containing elliptical speech cues, the analysis showed a significant difference between these two conditions. This difference was small compared to the difference in the scores when both sets of sentences were presented in the clear.

Cochlear Implant Patients. The mean sentence repetition performance results for six of the 10 cochlear implant users for both normal and anomalous sentences are shown in Figure 7. The remaining four cochlear implant users, whose CNC scores were very low, were unable to carry out the repetition task at all and their data were not included in this analysis.

The results shown in Figure 7 were surprising. The six cochlear implant users scored better on sentences containing intact speech cues than sentences containing elliptical speech cues. We expected that cochlear implant users would show performance that was similar to data obtained from normal-hearing listeners under signal degradation; however, the patients' performance was more similar to normal-hearing listeners in the clear. Somehow, the cochlear implant patients were able to distinguish between the two forms of speech.

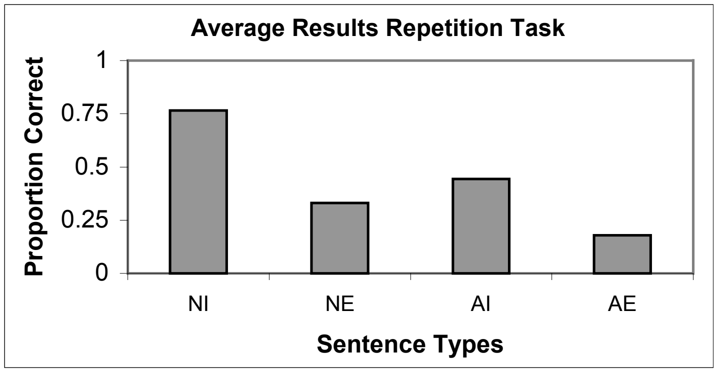


Figure 7. Average results from the repetition task for 6 cochlear implant patients.

The results were analyzed using a 2x2 ANOVA. One factor was speech cues (intact vs. elliptical cues). The second factor was meaningfulness (normal vs. anomalous sentences). Significant main effects were observed for both speech cues (F (1,5) = 150.1, p<.001) and sentence types (F (1,5) = 221.8, p<.001) suggesting that the cochlear implant patients were able to discriminate between intact and elliptical speech in this sentence repetition task.

The individual scores for the six cochlear implant patients that were able to complete the repetition task are shown in Figure 8 in order of increasing CNC score. An implant patient's CNC score is shown next to the corresponding number in the legends of each figure. Each of the patients scored better on sentences containing intact speech cues than the sentences containing elliptical speech cues. Performance on the sentence repetition task coincides with the CNC score for almost all of the cochlear implant users for each sentence type.

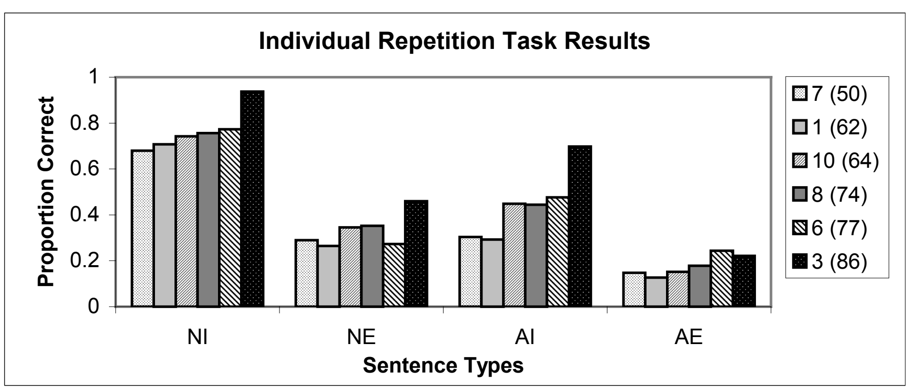


Figure 8. Individual results from the repetition task for the 6 cochlear implant patients.

General Discussion

Despite difficulties in perceiving fine phonetic contrasts in speech, such as place of articulation in consonants, many cochlear implant users are able to comprehend fluent speech and recover the talkers' intended linguistic message. What does the speech sound like for users of cochlear implants? How do patients with cochlear implants manage to comprehend spoken language despite receiving highly degraded sensory information? The results obtained from the two experiments reported in this paper provide some interesting new insights into the underlying perceptual processes and suggest some possibilities for intervention and oral rehabilitation in adult patients in the weeks and months immediately after they receive their cochlear implant.

Same-Different Discrimination Task

The first condition in this study used a same-different discrimination task with pairs of sentences that had either intact place of articulation cues or elliptical speech cues. The findings from normal-hearing listeners replicated the informal observations made by Miller and Nicely (1955) based on confusion data obtained in noise or under lo-pass filtering. Miller and Nicely originally suggested that elliptical speech, which is impoverished with respect to specifying place of articulation, may not be perceived as deficient under degraded listening conditions because these conditions "reinstate" or "reproduce" the original conditions that produced the signal degradation.

The present findings also replicate the earlier results reported by Quillet et al. (1998), which employed a similar same-different task and the results from Herman and Pisoni (2000) who used the same procedures with an exceptionally good patient with a cochlear implant. In both of these earlier experiments and the present study, normal-hearing listeners labeled pairs of sentences that were lexically identical but different because one contained intact speech cues and the other contained elliptical speech cues, as the "same" a majority of the time when heard under degraded conditions. When pairs of sentences were presented in the clear, normal-hearing listeners could easily discriminate intact speech from elliptical speech. However, when presented under degraded listening conditions such as masking noise, place of articulation information became less reliable and listeners tended to be unable to discriminate intact and elliptical speech. Listeners responded "same" almost all of the time based on broad phonetic coding of the stimulus input.

The cochlear implant patients that we studied here were also not able to distinguish intact speech cues from elliptical speech cues most of the time. This pattern was shown in Figure 2. The present results with a group of cochlear implant patients replicate the findings obtained in the case study reported by Herman and Pisoni (2000) who used only one cochlear implant patient. The present results suggest that phonetic contrasts such as place of articulation in consonants may not be perceived very well by cochlear implant users despite the fact that they can recognize spoken words and understand sentences. From an examination of the individual data, we observed a great deal of variability in performance across the 10 cochlear implant users. The "fair" users with CNC scores in the range from 30-40 had a great deal of trouble even discriminating sentences that were lexically different. In contrast, the best implant user who had a CNC score of 86 could consistently discriminate between intact and elliptical speech most of the time. The remaining users who fell somewhere between these two extremes appeared to produce results that were similar to normal-hearing listeners under degraded listening conditions.

The overall pattern of results observed in the same-different discrimination task by normal-hearing listeners under degraded listening conditions and by cochlear implant users in the quiet were similar to each other despite some small differences in procedure. Both groups appeared to rely on coarse phonetic coding in which place of articulation differences were no longer perceptually prominent. Both groups perceived the intact version and the elliptical versions of a sentence as the “same” under these presentation conditions. The present findings demonstrate the use of broad perceptual equivalence classes in speech perception when only partial stimulus information is present in the signal as originally suggested by Miller and Nicely (1955). This pattern of results resembles the earlier results reported by Herman and Pisoni (2000) and the preliminary findings from Quillet et al. (1998) for normal-hearing listeners under degraded listening conditions. Despite difficulty making fine lexical discriminations between spoken words, hearing-impaired individuals who use cochlear implants are able to use coarse-coding strategies to support spoken word recognition and speech perception. Many of these patients are able to derive substantial benefits from their cochlear implants even when they receive degraded and somewhat impoverished acoustic-phonetic information.

Sentence Repetition Task

The second condition in this study used an immediate repetition task with sentences that had either intact place of articulation cues or elliptical speech cues. These sentences were presented either in the clear or in noise. The findings from the normal-hearing listeners failed to support our initial predictions that repetition performance for intact speech and elliptical speech under degraded listening conditions would be equivalent. While the results for intact and elliptical speech in noise look very similar on average, statistical analysis revealed a small but significant difference between the two conditions. Thus, sentence repetition performance for normal-hearing listeners under degraded conditions for the intact speech was higher than performance of elliptical speech.

The results from the sentence repetition task were not consistent with Miller and Nicely’s (1955) predictions or the earlier findings of Quillet et al. (1998) using a transcription task. However, the results replicate the more recent findings reported by Herman and Pisoni (2000) who reported that overall transcription performance under degraded conditions was slightly better for intact speech than elliptical speech. Their results suggest that some phonetic cues to place of articulation can be extracted from these stimuli. Our findings suggest that some additional phonetic cues to place of articulation are present in natural speech and these are not eliminated by using elliptical speech.

The performance of the cochlear implant patients on normal and anomalous sentences with intact and elliptical speech cues presented in the clear and in noise followed the same basic pattern observed in the normal-hearing listeners under degraded listening conditions. The pattern of results observed with these listeners also failed to support our original predictions. Similar to the results from Herman and Pisoni’s single cochlear implant user, each cochlear implant user’s repetition performance, regardless of their CNC word score, showed worse repetition performance for elliptical speech relative to speech with intact place of articulation. As shown in the individual results, the recognition of both intact and elliptical speech increased with increases in the CNC word recognition scores, although both conditions were quite different. This pattern suggests that like normal-hearing listeners who were able to perceive speech under degraded conditions, cochlear implant users were also able to perceive and make use of some weak phonetic cues even with presumed loss of information about place of articulation.

Although failing to support our original predictions, the patients with cochlear implants did show good transcription performance for the normal sentences compared to anomalous sentences, despite the impoverished phonetic input from their implants. The earlier observations of Shipman and Zue (1982) and Zue and Huttenlocher (1983) concerning the powerful sound sequencing constraints in English and their role in spoken word recognition are consistent with the present findings obtained from cochlear implant users in both tasks. The patients are clearly able to make extremely good use of the minimal speech cues available to them in order to reduce the search space and to permit reliable lexical access and sentence comprehension. In many ways, these are impressive accomplishments given the highly degraded nature of the acoustic-phonetic input these patients receive.

The use of elliptical speech as a research tool to study sentence perception may lead not only to a better understanding of which speech sounds patients can discriminate after cochlear implantation and which ones they cannot, but these new stimulus materials may also help us to create better methods for training awareness of difficult phonological contrasts in patients with cochlear implants and help them deal with speech and spoken language in more efficient and optimal ways as they gain more experience with their cochlear implant and the nature of the transformation it imposes on speech signals.

In summary, the present investigation of normal-hearing listeners and a small group of cochlear implant patients provided an interesting pattern of results that differed from our original predictions. In the sentence discrimination task, both normal-hearing listeners and cochlear implant users were unable to discriminate between intact and elliptical speech suggesting the use of broad equivalent classes in speech perception and spoken word recognition. However, in the sentence repetition task, both groups were able to discriminate between intact and elliptical speech cues. This pattern of results suggests that some acoustic-phonetic cues to place of articulation can be perceived reliably. Although we observed differences in the two tasks, the present findings suggest that patients with cochlear implants perceive speech using broad perceptual equivalence classes and they appear to make use of less detailed phonetic categories than normal-hearing listeners routinely do under the same listening conditions.

References

Egan, J.P. (1948). Articulation testing methods. *Laryngoscope, 58,* 955-991.

Herman, R. & Pisoni, D.B. (2000). Perception of “elliptical speech” by an adult hearing-impaired listener with a cochlear implant: some preliminary findings on coarse-coding in speech perception. In *Research on Spoken Language Processing Progress Report No. 24* (pp. 87-112) Bloomington, IN: Speech Research Laboratory, Indiana University.

IEEE. (1969). IEEE recommended practice for speech quality measurements. *IEEE Report No. 297*.

Miller, G.A. (1956). The perception of speech. In M. Halle (Ed.) *For Roman Jakobson* (pp. 938-1062.) The Hague: Mouton.

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

Quillet, C., Wright, R. & Pisoni, D.B. (1998). Perception of “place-degraded speech” by normal-hearing listeners: Some preliminary findings. In *Research on Spoken Language Processing Progress Report No. 22* (pp. 354-375). Bloomington, IN: Speech Research Laboratory, Indiana University.

Shipman, D.W., & Zue, V.W. (1982). Properties of large lexicons: Implications for advanced isolated word recognition systems. In *ICASSP 82 Proceedings*, *May 3-5, 1982*.

Zue, V.W. & Huttenlocher, D.P. (1983). Computer recognition of isolated words from large vocabularies. *ICASSP 83: Proceedings, April 14-16, 1983, Sheraton-Boston Hotel, Boston, Massachusetts, IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 121-125). New York: The Institute.