A Preliminary Comparison of Performance between Patients Fit with the CII Bionic Ear™ and Patients Fit with the Nucleus 3G System

Michael F. Dorman, Anthony Spahr
Arizona State University
Tempe, Arizona 85158

Karen I. Kirk
Indiana University School of Medicine, Indianapolis, Indiana 46202

INTRODUCTION

The three cochlear implants available in the U.S. differ in terms of signal processing strategies and the hardware that implements those strategies. It is likely that some of these differences lead to differences in patient performance. In the preliminary results reported here, the outcome of an experiment in which the perception of speech, voices, and music was assessed in patients fit with the CII Bionic Ear™ with HiResolution sound processing (Advanced Bionics Corporation) and in patients fit with the Nucleus 3 system (Cochlear Corporation) are described. In the near future, results for patients fit with the Med El Tempo+ device will be added. The aim was to look for differences in performance as a means to determine which aspects of implant signal processing and design lead to improved implant benefit.

The study used a novel design for comparing performance between devices. Patients were matched first by their CNC word scores in quiet, and then by age. For example, a CII patient with a 70% CNC score was matched with a 3G patient who scored 70% correct, and a CII patient with a 45% CNC score was matched with a 3G patient with a 45% correct score. The matched groups then were tested and compared using stimuli presented in quiet and in noise. At issue was whether the patients differed in performance on tests of detailed spectral resolution, detailed temporal resolution, speech understanding in noise, and speech perception at low sound levels.

SUBJECTS

Patients were recruited by letter from implant centers in the United States and Canada. Patients had to score greater than 45% correct on CNC words to be included in the study. Patients with CNC scores lower than 45% were excluded from the study because they showed a floor effect when tested with difficult sentence material (AzBio sentences) at +10 and at +5 dB SNR.

A total of 19 CII patients and 29 3G patients have been tested and have met the study criteria. From these two groups, 13 matched pairs of subjects have been identified whose CNC scores matched within ± 4 percentage points (i.e., ± 2 items on one 50-word CNC list). The CII group mean score was 64% correct, and the 3G group mean score was 65% correct. The mean ages for subjects in the two groups did not differ (CII mean = 57 years, 3G mean = 58 years). Subjects were not matched for any other demographic characteristics.

TEST MATERIALS

New Test Materials

To ensure that patients from one group did not have more experience with the test materials than patients from the other group, a new battery of tests was created.

AzBio sentences. The AzBio sentence test consists of 500 recorded sentences ranging in length from six to ten words. A total of five talkers (two male and three female) were used. All sentences were normalized to be of approximately equal SPL. The sentences then were processed using a five-channel cochlear-implant simulation and presented to ten normal-hearing subjects for identification. Mean percent correct scores then were calculated for each of the 500 sentences. Based upon those intelligibility scores, nine lists of 40 sentences each were constructed using sentences produced by four talkers (two male and two female). The mean intelligibility of the lists (for normal hearing subjects listening to the five-channel simulation) was 89% correct and differed by less than 1% across lists.

Implant subjects were asked to repeat back the sentences and were encouraged to guess when unsure. All individual sentences and lists were scored as words correct and an overall percent correct score was computed.

Talker discrimination. Single-word stimuli were selected from a digital database developed at the Speech Research Laboratory at Indiana University, Bloomington. The stimulus set consisted of 108 words produced by five males and five females. Implant subjects were presented with pairs of words. Within each condition, half of the pairings were produced by the same talker, and half were produced by different talkers. The words in each pairing were always different. For example, Male Talker 1 might say "ball" and Male Talker 2 might say "brush." Across the different-talker pairs, each talker was paired with every other talker an equal number of times. Participants responded "same" or "different" by pressing one of two buttons. Responses were scored as the percent of correct responses (Kirk et al., 2002).

Melody recognition. Each subject selected five familiar melodies from a list of 33 simple melodies. Each melody consisted of 16 equal-duration notes synthesized with MIDI software that used samples of a grand piano (Hartmann and Johnson, 1991). The frequencies of the notes ranged from 277 Hz to 622 Hz. The average note was concert A (440 Hz) ± 1 semitone. The melodies were created without distinctive rhythmic information. After presentation of a melody, subjects responded by pressing a button from a list containing his or her five pre-selected melodies.

Vowel recognition without duration cues. Thirteen vowel stimuli were created in /bVt/ format using KLATT software ("bait, Bart, bat, beet, Bert, bet, bit, bite, boat, boot, bought, bout, but"). Vowel duration was synthesized to be brief (90 msec) and equal across all 13 vowels so that vowel length would not serve as a cue for identification (Dorman et al., 1989). There were five repetitions of each stimulus. The order of the stimuli was randomized in the test list.

Consonants in /e/ environment. Twenty consonants were recorded in /eCe/ format, e.g., 'a bay, 'a day', a gay', etc. A single male talker produced five variations of each token. The pitch and vocalic portion of each token were varied intentionally. The order of items was randomized in the test list.

Standard Test Materials

CNC words. One 50-item CNC word list was presented to each subject.

Hearing In Noise Test (HINT). All 250 HINT sentences were presented in random order in quiet (Nilsson et al., 1994).

City University of New York (CUNY) Sentences. CUNY sentences were used in quiet and at +10 dB SNR. Two lists, for a total of 24 sentences, were used in each condition. All lists were taken from the Cochlear Corporation Investigational Test Battery CD (Boothroyd et al., 1988).

Noise

The noise was four-talker babble. The noise started 100 ms before the onset of the signal and ended 100 ms after the end of the signal.

PROCEDURES

Before testing, subjects were instructed to adjust their processors to the settings they most commonly used. They then were evaluated using the test battery. The patients received practice on each type of test material before the test began. All signals were delivered through a loudspeaker located directly in front of the listener. The same test order was used for all subjects: (1) talker discrimination, (2) CNC words, (3) synthetic vowels, (4) consonants, (5) CUNY and AzBio sentences at 74 dB SPL in quiet, (6) CUNY and AzBio sentences at 74 dB SPL with +10 SNR, (7) AzBio sentences at 74 dB SPL with +5 SNR, (8) AzBio sentences at 64 SPL in quiet, (9) AzBio sentences at 54 dB SPL in quiet, (10) 250 HINT sentences in quiet and (11) melodies.

In addition, a new metric, termed "robustness," was developed to evaluate performance in the two difficult listening conditions (in noise, at low SPLs). Robustness was computed by averaging the scores for the two difficult listening situations (74 dB SPL @ +10 SNR and 54 dB SPL in quiet) and then dividing the average score by the score obtained at 74 dB SPL in quiet. That number was multiplied by 100. A high score on the robustness index indicates little drop in performance between the 74 dB quiet condition and the two difficult listening situations. A low score indicates a large drop in performance between the 74 dB quiet condition and one or both of the difficult listening situations.

RESULTS

The results are shown in Figures 1-5. Figure 1 shows that the mean performance on CUNY sentences was substantially higher than on the AzBio sentences, a trend that also was observed for the HINT sentences. Moreover, many patients reached the test ceiling for the CUNY and HINT sentences, thereby indicating that these measures are no longer adequate to detect differences in device performance. In contrast, the AzBio sentences are more difficult and thus more sensitive to device differences than the CUNY and HINT sentence materials.

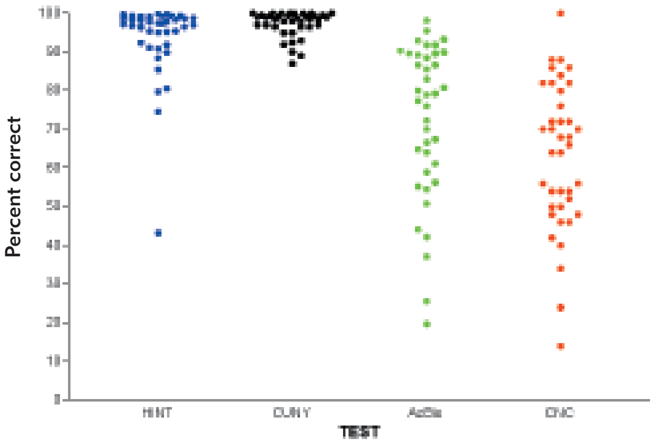

Figure 1. Individual scores on HINT sentences, CUNY sentences, AzBio sentences and CNC words in quiet.

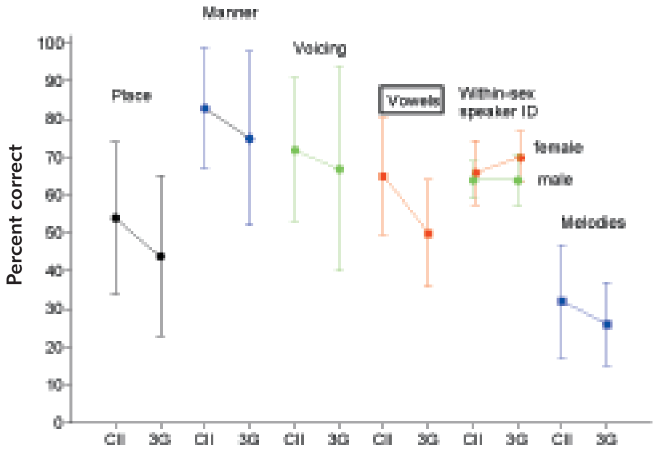

Figure 2. Performance on tests of spectral and temporal resolution for patients matched on CNC scores and fit with the CII Bionic Ear™ with HiResolution sound processing (CII) or the Nucleus 3G system (3G). Differences between groups were significant only for vowel recognition.

Significant differences in performance ($p < 0.05$) were found between the CII and 3G subjects in four test conditions. As shown in Figure 2, CII patients recognized brief synthetic vowels with better accuracy than 3G patients (CII average = 65%; 3G average = 49%).

As shown in Figure 3, CII patients recognized AzBio sentences at +10 dB SNR better than 3G patients (CII average = 49%; 3G average = 32%) and CII patients recognized AzBio sentences at +5 dB SNR better than 3G patients (CII average = 34%; 3G average = 14%).

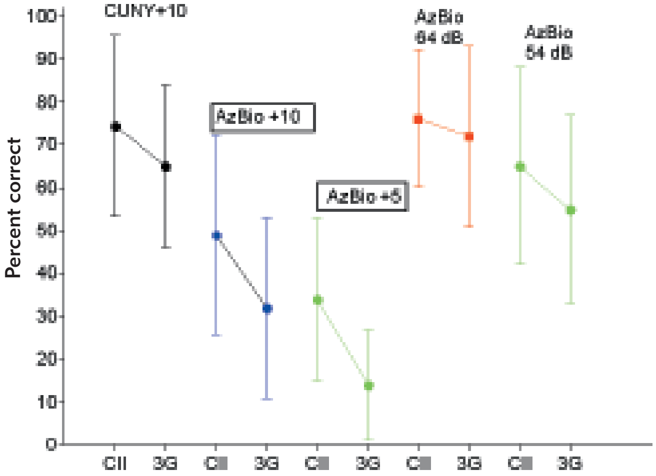

Figure 3. Performance in difficult listening conditions for patients matched on CNC scores and fit with the CII Bionic Ear™ with HiResolution sound processing (CII) or the Nucleus 3G system (3G). Differences between groups were significant only for the AzBio sentences in the +10 and +5 dB signal-to-noise conditions.

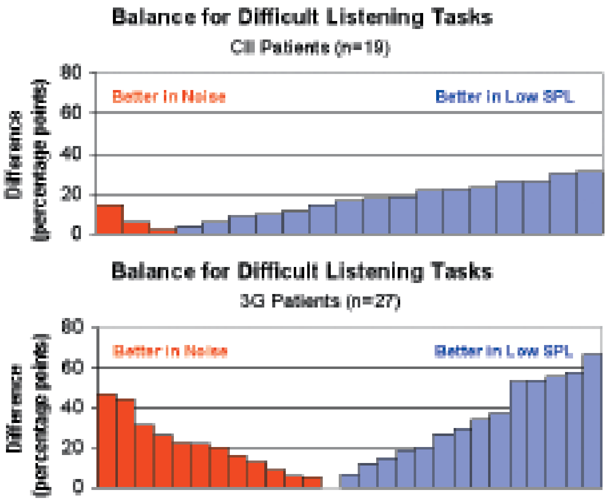

Figure 4. Percentage point differences between scores in noise and at a low input level for patients fit with the 3G and CII systems. The bar for each patient represents the percentage point difference between scores in noise (+10 dB SNR) and low input level (54 dB SPL). If the score in noise was better than the score at low input level then the bar was colored red and placed on the left side of the figure. If the score for the low input level was better than the score in noise then the bar was colored blue and placed on the right side of the figure.

As shown in Figure 4, CII patients tended to have relatively equal levels of performance in noise and in the low signal level condition. The majority of patients had a 20 percentage point or less difference between the two conditions. The 3G patients evidenced two different patterns of performance. Either the score in noise was much better than the score at the low level or the score at the low level was much better than the score in noise.

The robustness metric quantifies performance in difficult listening situations relative to performance in quiet. As shown in Figure 5, CII patients had significantly higher robustness scores ($p < 0.05$) than the 3G users (CII average = 73%; 3G average = 61%).

For all other tests and conditions, the differences in performance between groups were not significant.

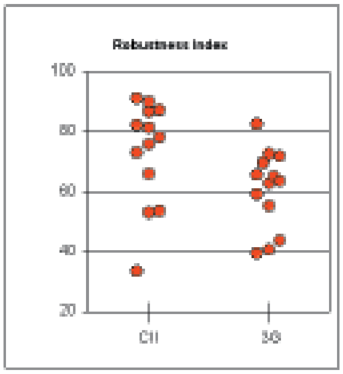

Figure 5. Performance in difficult listening situations as a function of performance in quiet. A score of 100 indicates no drop in performance relative to quiet 74 dB SPL for performance at +10 dB SNR and at 54 dB SPL.

DISCUSSION

The long-term goal of this project is to identify aspects of cochlear-implant signal processing and hardware design that result in improvements in patient performance. The first step toward this goal is to determine whether differences exist in the performance of patients fit with different devices that are currently in use. It is encouraging to find that there are, in fact, differences in performance between groups of CII and 3G patients matched for CNC word scores and for age.

CII patients scored higher than 3G patients on the task of identifying synthetic vowels in a "bVt" context. Because the vowels were relatively brief and did not differ in duration, the task likely was one of spectral resolution. As a first guess, the differences in performance between the groups may result from the larger number of output channels in the CII device (16 for the CII patients and either 8- or 12-of-20 for the 3G patients).

CII patients also scored higher than 3G patients in both the +10 and +5 dB SNR conditions on the difficult AzBio sentences. The difference in number of output channels may be responsible for this outcome since other experiments have shown that performance in noise is related directly to the number of channels used to code the stimulus (Dorman et al., 1998, Fu et al., 1998).

CII patients tended to have relatively equal scores (20% or less difference) in noise and in the low signal level condition. In contrast, 3G patients evidenced a greater difference between conditions. A good balance between performance in noise and at low signal levels would seem to be a desirable characteristic of an implant.

The measure of balance described above does not take into account the level of performance in quiet, i.e., the patients with a small difference score could do equally well or equally poorly in the difficult listening situations. The measure of robustness takes this into account and on this measure CII patients scored higher than 3G patients. Thus, the drop in performance in the two difficult listening conditions was less for the CII patients than the 3G patients.

CONCLUSIONS

Given the nature of this research design, it is encouraging that there are differences in performance between patients using the two implant systems. Thus, the hardware and software differences that underlie the differences in performance can be probed. However, at this early stage of this project, the reasons for the differences in performance are not clear. Additional information will be obtained as more subject pairs are added to the groups (to ensure that group differences are real) and when more Med El patients have been tested.

As a caveat, the results obtained in these experiments are relevant only to patients who understand speech relatively well with their cochlear implants, i.e., those with 40-45% CNC scores or better. Patients with CNC scores lower than 40-45% were not included because, when tested with difficult sentence material, their scores in noise tend to exhibit a floor effect, thereby precluding the ability to accurately measure a drop in performance. Consequently, the results do not represent a comparison of overall patient performance across devices.

Acknowledgment

This research was supported by NIDCD grant RO1 00654-13 and by contributions from Advanced Bionics Corporation, Cochlear Americas, and Med El Corporation.

References

Boothroyd, A., Hnath-Chisolm, T., Hanin, L., Kishon-Rabin L. (1988). Voice fundamental frequency as an auditory supplement to the speechreading of sentences. *Ear & Hearing*, 9(6), 306-12.

Dorman, M., Dankowski K., McCandless, G., Smith L. (1989). Identification of synthetic vowels by patients using the Symbion multichannel cochlear implant. *Ear & Hearing*, 10(1), 40-43.

Dorman, M., Loizou, P., Fitzke, J. and Tu, Z. (1998) The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *Journal of the Acoustical Society of America*, 104 (6), 3583-3585.

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

Hartmann, W.M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception*, 9, 155-184.

Kirk, K.I., Houston, D.M., Pisoni, D.B., Sprunger, A.B., & Kim-Lee, Y. (2002). Talker discrimination and spoken word recognition by adults with cochlear implants. Poster presented at the 2002 Mid-Winter Meeting for the Association for Research in Otolaryngology, St. Petersburg Beach, Florida.

Nillson, M., Soli, S., Sullivan, J. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, 95(2), 1085-99.

June 2003

The authors are solely responsible for study design, data analyses, and interpretation of results. Manufacturer support does not imply endorsement of or agreement with study conclusions.