Title: The recognition of monosyllabic words by cochlear-implant patients and by normal-hearing subjects listening to words processed through cochlear implant signal processing strategies.

Running head: Word recognition

Authors: Michael F. Dorman, Ph.D., Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-0102 and Division of Otolaryngology, University of Utah Health Sciences Center; Philip C. Loizou, Ph.D., Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083-0688; Jeanette Fitzke, M.A., Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-0102 and Zhemin Tu, M.S., Department of Applied Science, University of Arkansas, Little Rock, 72204-1099

Grant support: This work was supported by R01 DC-000654-07 from the NIDCD.

Reprint requests and correspondence to:

Michael F. Dorman, Ph.D., Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-0102. Telephone 602-965-3345; fax 602 965-0965

Email mdorman@asu.edu

Introduction

The aim of a cochlear implant is to generate, using artificial electrical stimulation, a pattern of neural activity that conveys to a patient the information in speech signals, music or environmental sounds. Given recent advances in signal processing and in electrode design, it is appropriate to ask how well the current generation of implants achieves the goal of transmitting the relevant information in the speech signal. One way to answer this question is to process signals in the manner of a cochlear implant signal processor and present the signals acoustically, using noise bands[1] or sine wave stimulation[2], to normal-hearing listeners. The performance of the normal-hearing listeners establishes a benchmark for how well implanted patients could perform *if* electrode arrays were able to reproduce, by artificial electrical stimulation, the stimulation produced by auditory stimulation of the cochlea, and *if* patients possessed neural structures capable of responding to the electrical stimulation. In this report, we describe the performance of normal-hearing subjects, tested with monosyllabic words processed through fixed-channel, CIS-like signal processing strategies and spectral maximum, SPEAK-like strategies and the performance of patients who use cochlear implants with CIS and SPEAK strategies.

Method

Subjects. Sixteen normal-hearing listeners were tested with NU-6 words processed through fixed-channel processing algorithms. Nine listeners were tested with words processed in the manner of a SPEAK-like algorithm The NU-6 scores for 25 patients fit with the 6-channel, MED EL CIS-LINK processor were obtained from records at the University of Iowa and the University of Utah. The patients had used their processors for periods ranging from 3 months to 3 years. The scores from 60 patients using the 8-channel MED EL COMBI-40 processor were provided by MED EL Corporation. The patients had used their processors for 1 year. The scores of 60 patients using the 8-channel CLARION 1.2 processor were provided by Advanced Bionics Corporation. The patients had used their processors for 6 months. The scores of 59 patients using the NUCLEUS 22 device and the SPEAK strategy were accumulated from records at the University of Michigan, the House Ear Institute and the University of Arkansas at Little Rock. The patients had used their processors for periods ranging from 6 months to 2 years. The scores of 66 patients using the Nucleus CI24M device and SPEAK strategy were provided by Cochlear Corporation. The patients had used their devices for 6 months.

Signal processing. To produce words for the normal-hearing listeners, signals were first directed to a pre-emphasis filter and then bandpassed into 4, 6, 8 and 12 frequency bands using sixth-order Butterworth filters. The spacing of the frequency bands was logarithmic for the 4, 6 and 8 channel processors. For the 12 channel processor, linear spacing of bands was used up to 1 kHz and logarithmic spacing thereafter. The envelope of the signal was extracted by full-wave rectification and low-pass filtering at 400 Hz. Sinusoids were generated with amplitudes equal to the root-mean-square (rms) energy of the envelopes (computed every 4 msec) and frequencies equal to the center frequencies of the bandpass filters.

For the SPEAK-like, or "n of m"strategy, the number of "m" channels was fixed at 20 and every 4 msec the 2, 4 or 6 "n" channels with the maximum energy were identified and output as sine waves at the center frequency of the filter band.

Procedure. The normal-hearing subjects were tested with 50 words from one of the NU6 word lists in each channel condition. Practice for each condition consisted of 50 words processed in the manner of that condition. The implanted subjects in the United States were tested with one of the 50 word NU6 lists. The subjects tested in Europe were tested with monosyllabic word lists appropriate for their language. Responses were scored in terms of words and phonemes correct.

Results

The results for the normal-hearing subjects listening to signals processed by fixed-channel strategies are shown as the filled squares in each of the four panels of Figure 1. The mean word scores for processors with 4, 6, 8 and 12 channels were 48 % correct, 71 % correct, 87 % correct and 89 % correct, respectively. The mean phoneme scores were 67 % correct, 85 % correct, 95 % correct and 96 % correct, respectively.

The performance of the 6 channel CIS-LINK patients is shown in Figure 1a. Five of the 25 patients achieved scores within or just outside of the standard deviation for the performance of the normal-hearing listeners. The median score was 38 percent correct and the interquartile range was 27 to 70 percent correct.

The performance of the CLARION 1.2 patients is shown in Figure 1b. Seven of the 60 patients achieved scores within or near the standard deviation for the performance of the normal-hearing subjects listening to words processed into 8 channels. The median score was 41 percent correct and the interquartile range was 15 to 60 percent correct.

The performance of the MED EL COMBI-40 patients is shown in Figure 1c. Five of the 60 patients achieved scores within or near the standard deviation for the performance of the normal-hearing subjects listening to words processed into 8 channels. The median score was 50 % correct and the interquartile range was 32 to 63 percent correct.

The mean scores for normal-hearing subjects when listening to signals processed by 2-, 4- and 6-of-20 SPEAK-type processors were 73, 87 and 90 % correct, respectively. The score for the 6-of-20 strategy, the one most similar to the SPEAK strategy, is shown in Figure 1d along with the scores for fixed-channel processors and the scores for Nucleus 22- SPEAK patients and CI24M-SPEAK patients. None of the Nucleus 22 subjects achieved scores within the standard deviation of scores of normal-hearing subjects listening to either an 8 channel processor or a 6-of-20 processor. The median score was 28 percent correct and the interquartile range was 12 to 44 percent correct. Two of the CI24M patients achieved scores within the standard deviation of scores for normal hearing subjects listening to an 8-channel processor. The median score was 42 percent correct and the interquartile range was 25 to 56 percent correct. .

Discussion. A number of patients using devices with eight or more channels achieved scores which were within the range of scores achieved by normal-hearing subjects listening to words processed into eight channels. This outcome is particularly impressive because the mean score for the normal hearing listeners -- 87 % -- is at such a high level. However, because the normal-hearing listeners had very limited experience with signals processed into a small number of channels, the estimate of potential performance may be too conservative. If the performance of normal-hearing listeners improved only a little with extended practice, e.g., 10 %,, then the scores of the best performing patients, with only two exceptions, would fall within the range of performance of normal-hearing subjects listening to signals processed into 6 channels. Given this possibility, we conclude that electrical stimulation of the cochlea is capable of reproducing with high fidelity the acoustic information in six to eight channels.

On the other hand, our results indicate that, on average, patients achieve scores which are within the range of scores achieved by normal-hearing subjects listening to speech processed into four channels. This outcome indicates a need for significant improvements in cochlear implants.

It is important to note that the two 8-channel devices employ different electrode arrays and patients fit with each of the arrays were able to achieve scores within the range of the normal-hearing listeners. The performance of the 6-channel CIS-link patients is of interest, in this light, because the 6 monopolar electrodes are spaced 4 mm apart and the most basal electrodes stimulate a range of frequencies far outside the normal range of frequencies in the speech signal. Cortical recognition routines for speech must be extremely flexible to accommodate the large distortions in the frequency representation of speech engendered by electrical stimulation of the cochlea.

The data reported here for implant patients can not be used to compare the efficacy of one implant device versus another. The population samples were not matched for variables, such as length of device use and length of deafness, which are known to affect performance. Another variable, the level of preoperative speech reception, was not controlled and appears to account for the higher mean scores of patients using the Nucleus CI24M than of patients using the Nucleus 22. Rubinstein et al.[3] report that, in the Iowa sample of implant patients, Nucleus 22 patients and CI24M patients implanted in the 1990s had significantly better preimplant speech recognition scores than Nucleus 22 patients implanted in the 1980s. Test scores for the later implanted Nucleus 22 patients and Nucleus CI24M patients were similar and both were higher than the scores for the earlier implanted Nucleus 22 patients. Thus, the improved scores for the CI24M patients appears to be related to a patient variable, perhaps the number and distribution of spiral ganglion cells, rather than a processor variable.

References

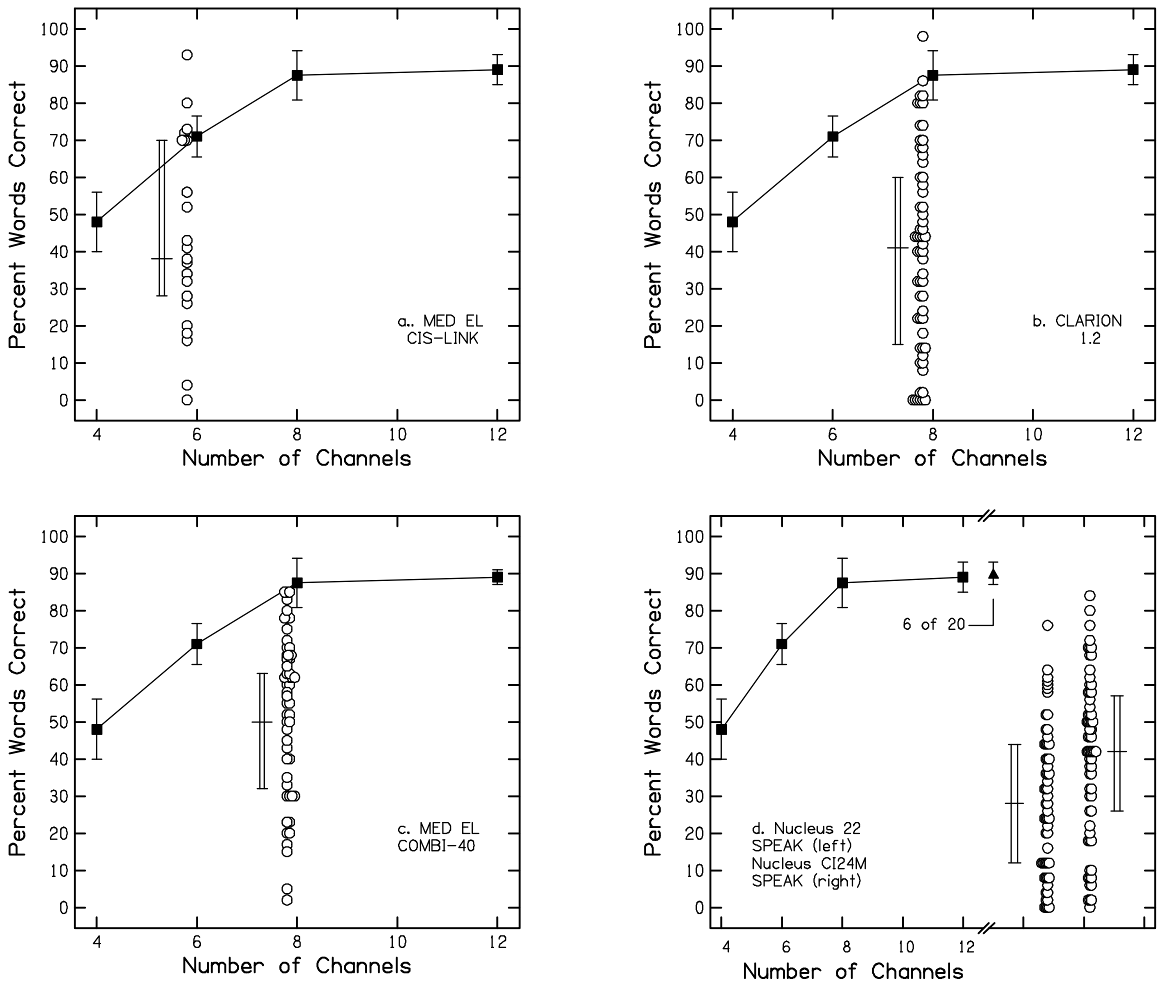
1. Dorman M., Loizou P., and Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, J. Acoust. Soc. Am. 1997, 102(4): 2403-2411.

2.Shannon R., Zeng F-G., Kamath V., Wygonski J. and Ekelid M. Speech recognition with primarily temporal cues, Science 1995, 270, 303-304.

3. Rubinstein J., Parkinson W., Tyler R, and Gantz B. Residual speech recognition and cochlear implant performance: Effects of implantation criteria, Am. J. Oto in press.

Figure caption

Figure 1. Percent correct word recognition by normal-hearing listeners (solid symbols) and by patients (open circles) fit with (a) the MED EL CIS-LINK processor, (b) the Advanced Bionics CLARION 1.2 processor, (c) the MED EL COMBI-40 processor and (d) the Cochlear Corporation Nucleus 22 (on left) and CI24M device (on right). Each open circle represents a patient's score. The median and interquartile range of scores for the implant patients are indicated by the rectangles next to the patient data.


a.. MED EL
CIS-LINK
b. CLARION
1.2
c. MED EL
COMBI-40
d. Nucleus 22
SPEAK (left)
Nucleus CI24M
SPEAK (right)
6 of 20
Percent Words Correct
Number of Channels