Ear & Hearing: Volume 19(2) April 1998 pp 162-166

The Identification of Consonants and Vowels by Cochlear Implant Patients Using a 6-Channel Continuous Interleaved Sampling Processor and by Normal-Hearing Subjects Using Simulations of Processors with Two to Nine Channels

Dorman, Michael F.; Loizou, Philipos C.

Arizona State University (M.F.D.), Tempe, Arizona 85287-0102, University of Utah Health Sciences Center (M.F.D.), Salt Lake City, Utah; and University of Arkansas at Little Rock (P.C.L.), Little Rock, Arkansas.

Address for correspondence: Michael F. Dorman, Ph.D., Department of Speech and Hearing Science, Arizona State University, Tempe, AZ 85287-0102.

Received December 5, 1996; accepted November 5, 1997

Abstract TOP

Objective: To compare the vowel and consonant identification ability of cochlear implant patients using a 6-channel continuous interleaved sampling (CIS) processor and of normal-hearing subjects using simulations of processors with two to nine channels.

Design: Subjects, 10 normal-hearing listeners and seven cochlear implant patients, were presented synthetic vowels in /bVt/ context, natural vowels produced by men, women, and girls in /hVd/ context, and consonants in /aCa/ context for identification. Stimuli for the normal-hearing subjects were pre-processed through simulations of implant processors with two to nine channels and were output as the sum of sinusoids at the center frequencies of the analysis filters.

Results: Five implant patients' scores fell within the range of normal performance with a 6-channel processor when the patients were tested with synthetic vowels. Four patients' scores fell within the range of normal with a 6-channel processor when the patients were tested with multitalker vowels. Five patients' scores fell within the range of normal for a 6-channel processor for the consonant feature "place of articulation."

Conclusion: Signal processing technology for cochlear implants has matured sufficiently to allow some patients who use CIS processors and a small number of monopolar electrodes to achieve scores on tests of speech identification that are within the range of scores established by normal-hearing subjects listening to speech processed through a small number of channels.

Article Outline

A commonly asked question about cochlear implant patients is whether they can, or do, receive the same amount of information that a normal-hearing listener would receive if a normal-hearing listener were listening to signals processed in the manner of the implant. That is, can a cochlear implant patient receive the same amount of information from, for example, a 6-channel processor as a normal-hearing individual listening to speech through a simulation of the 6-channel processor? The answer to this question will indicate how well signal processing strategies, electrode arrays, and patients' surviving neural populations recreate the information in the speech signal.

To help answer the question posed above, **Dorman, Loizou, and Rainey (1997)** processed speech signals through software simulations of cochlear implant signal processors with two to nine channels and presented the signals, combinations of sine waves at the center frequencies of the analysis filters, to normal-hearing listeners for identification. The number of channels necessary to reach asymptotic performance varied as a function of the test material. The identification of consonant place of articulation, the identification of synthetic vowels that varied only in formant frequencies, and the identification of vowels produced by multiple talkers (men, women, girls, and boys) reached asymptote with eight channels of stimulation. The identification of sentences reached asymptote with five channels of stimulation. Thus, as **Shannon, Zeng, Kamath, Wygonski, and Ekelid (1995)** reported previously, high levels of identification accuracy can be obtained with a relatively small number of channels of stimulation.

High levels of identification accuracy also have been reported for cochlear implant patients who use signal processors with four, six, and eight channels of stimulation **(Brill et al., Reference Note 1; Tyler, Gantz, Woodworth, Parkinson, Lowder, and Schum, 1996; Wilson, Lawson, and Zerbi, 1995)**. These results raise the possibility that some implant patients can extract as much information from a small number of channels as normal-hearing listeners. In this report we compare the performance of individuals who use a 6-channel, Med El cochlear implant signal processor with the performance of the normal-hearing subjects tested by**Dorman et al. (1997)**. At issue is whether any of the implant patients can extract enough information from the speech signal to achieve scores on tests of speech recognition that are within the range of scores established by normal-hearing subjects listening to speech processed through six channels.

Method TOP

Subjects TOP

Ten normal-hearing students at Arizona State University, who ranged in age from 21 to 62 yr, and seven cochlear implant patients, who ranged in age from 37 to 73 yr, participated in the experiment. The implant patients had used an Ineraid cochlear implant, which uses a monopolar electrode array, for at least 5 yr before being fitted with a research version of the Med El cochlear implant signal processor. The patients had used the research processor for a minimum duration of 1 mo and a maximum duration of 3 yr. The patients were chosen for participation on the basis of availability for testing. Biographical data on the patients are presented in**Table 1**.

Patient	Age (yr) at Detection of Hearing Loss	Age at which Hearing Aid Gave no Benefit	Age Fitted with Ineraid
S1	19	19	29
S2	23	48	51
S3	7	31	33
S4	20	46	63
S5	5	43	48
S6	1	27	30
S7	20	46	63

TABLE 1. Biographical data on implant patients. S5 cannot remember ever hearing in the ear that was implanted. Sound presented to the implanted ear is lateralized to the opposite ear.

Speech Materials TOP

The speech stimuli were synthetic vowels in /bVt/ context, naturally produced vowels in /hVd/ context, and naturally produced consonants in /aCa/ context. These stimuli were a subset of the stimuli used in**Dorman et al. (1997)**. They were selected for use in the present experiment on the grounds that normal-hearing listeners did not evidence a ceiling effect in performance when the stimuli were processed through a 6-channel processor.

The stimuli for the synthetic-vowel test were 13 vowels in /bVt/ format, i.e., "beet, bit, bait, bet, bat, bought, but, boot, Bert, Bart, bout, bite, boat" (see **Dorman, Dankowski, Smith, & McCandless, 1989**). The prototypes for the stimuli were created, using the KLATT algorithm, by "copy synthesis" of the author's productions of the /bVt/ stimuli. The prototypes then were altered to make vowel duration constant across stimuli. This step was taken to ensure that listeners' identification would be based on resolution of the signals' formant frequencies. Because the stimuli were based originally on copy synthesis, the stimuli were both relatively natural sounding and highly intelligible. Each stimulus was synthesized with five formants. Each stimulus was composed of a 5 msec/b/-burst, a 5 msec silent interval, 30 msec /b/-transitions, a 90 msec vocalic nucleus, 50 msec /t/-transitions, 80 msec of silence, and final/t/-burst/aspiration of 50 msec. A gradually falling F0 contour was used for all stimuli. Five tokens of each stimulus were created for the test sequence. The stimuli were grouped into five blocks with each stimulus appearing once in a block. Stimulus order within a block was randomized.

The stimuli for the natural-vowel test were the words "heed, hid, hayed, head, had, hod, hud, hood, hoed, who'd, heard," each produced by three men, three women, and three girls. The stimuli were drawn from a set created by**Hillenbrand, Getty, Clark, and Wheeler (1995)**. Each stimulus appeared once in the randomized test sequence.

The stimuli for the Iowa consonant test were 16 consonants in /aCa/ environment spoken by a single male speaker **(Tyler, Preece,& Tye-Murray, 1986)**. Five tokens of each stimulus were created for the test sequence. The stimuli were grouped into five blocks with each stimulus appearing once in a block. Stimulus order within a block was randomized.

All test materials were stored on computer disk and were output via custom software routines using MATLAB (The MathWorks, Inc.) software and a 16 bit D/A converter.

Signal Processing for Cochlear Implant Patients TOP

The signal processor was an implementation of the **Wilson, Finley, Lawson, Wolford, Eddington, and Rabinowitz (1991)** continuous interleaved sampling (CIS) processor fabricated at the University of Innsbruck **(Zierhofer, Peter, Bril, Pohl, Hochmair-Desoyer,& Hochmair, 1994)**. The processor was a 6-channel design with 6th order band-pass filters, 400 Hz first order smoother, and full wave rectification. Channel center frequencies were 393 Hz, 639 Hz, 1037 Hz, 1680 Hz, 2730 Hz, and 4440 Hz. The channels were of equal width on a logarithmic scale. Signals were pre-emphasized above 1200 Hz. Pulse duration and pulse rate were optimized for each patient.

Signal Processing for Normal-Hearing Listeners TOP

A software version of the Med El signal processor was implemented using the MATLAB signal-processing toolbox. After the operations of band-pass filtering, rectification, and smoothing, the root mean square (rms) energy of the envelope in each channel was computed every 4 msec and sinusoids were generated with amplitudes equal to the rms levels. The sinusoids were summed and presented to the listeners at a comfortable level through Sennheiser HMD 410 headphones.

Procedure TOP

Normal-hearing subjects were tested with signal processors having two to nine channels **(Dorman et al., 1997)**. Testing was performed in the order 9-channel processor to 2-channel processor, sequentially, to maximize subjects' familiarity with the stimuli and the sound of the processed stimuli before encountering the conditions with the fewest channels of information. For each condition subjects were given practice before the test sequence began. Practice consisted of two repetitions of each test item, followed by a randomized sequence of the stimuli, all with visual, correct-answer feedback.

Practice for the cochlear implant patients consisted of two repetitions of the test items with correct-answer feedback. The implant patients had been tested previously with the consonants and synthetic vowels (always without correct-answer feedback) before being tested in the present experiment. The patients had not been tested previously with the multitalker vowels.

Responses were collected with custom software using a computer display of response alternatives and a mouse as a response key. The subjects were allowed to use a "repeat" key as many times as they wished. Feedback of correct responses was not provided during the test.

Results TOP

The results are shown in **Figure 1a to 1c**, where the means and SDs of performance for normal-hearing listeners in the 2- to 9-channel conditions are replotted from **Dorman et al. (1997)** and where the performance of implanted patients is shown by symbols not connected by lines. As shown in **Figure 1a**, when the test material was synthetic vowels, normal-hearing subjects listening to six channels of stimulation achieved a mean score of 81% correct with an SD of 13. The scores of five implant patients fell within +/-1 SD of the mean for normal-hearing subjects. As shown in**Figure 1b**, when the test material was multitalker vowels, normal-hearing subjects listening to six channels of stimulation achieved a mean score of 80% correct with an SD of 6. The scores of four implant patients fell within +/-1 SD of the mean for normal-hearing subjects.

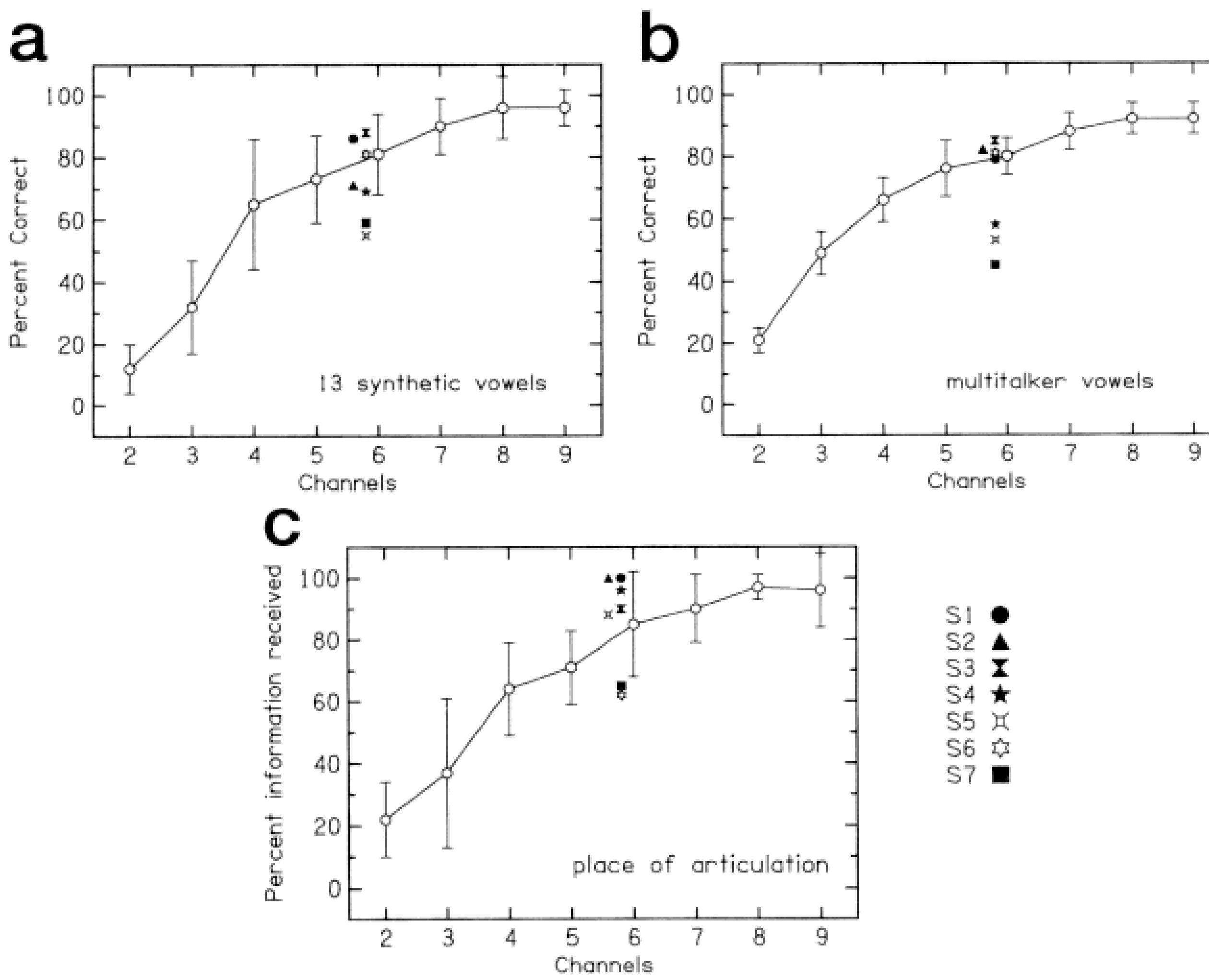

Figure 1. Test scores as a function of the number of channels of stimulation for normal-hearing listeners (open circles connected by solid lines) and for patients using a 6-channel continuous interleaved sampling processor (symbols not connected by lines).

An information transmission analysis for consonants was conducted using the features of **Miller and Nicely (1955)**. Only the data for place of articulation are plotted in **Figure 1c** because, for normal-hearing listeners in the 6-channel condition, only place of articulation was uninfluenced by a ceiling effect. The mean score for the normal-hearing listeners in the 6-channel condition was 85% correct with an SD of 17. The scores of five implant patients fell within +/-1 SD of the mean for normal-hearing subjects.

Discussion TOP

In the introduction we asked whether patients with cochlear implants can, or do, receive the same amount of information about speech as normal-hearing subjects listening to speech processed in the manner of the implant. The data from three tests of speech recognition suggest that some implant patients do achieve a level of speech recognition within the range established by normal-hearing listeners.

Interpretation of the data from the consonant identification task and the synthetic vowel identification task is constrained by two factors.* The first is that only one speaker (one real and one artificial) produced the utterances. The second is that the implanted patients had much more practice listening to the consonants and a little more practice listening to the synthetic vowels than the normal-hearing subjects. For example, all but one of the patients in the present study have participated in the research program at Research Triangle Institute(Lawson, Wilson, & Finley, 1993; Wilson et al., 1995) and have, in the week or weeks of testing, listened to the Iowa consonants a heroic number of times. Van Tassel, Greenfield, Logemann, and Nelson (1992)the Iowa consonants a heroic number of times. have argued that listeners can use idiosyncratic characteristics of single-talker stimuli, such as duration or pitch, to identify the stimuli. Because the implant patients had more practice with the consonants and synthetic vowels than the normal-hearing listeners, the scores of the implant patients may have been artificially inflated. In this light, the data from the multitalker vowels are especially probative because the implant patients had not been tested previously with these stimuli. The outcome that the scores of four implant patients fell within the range of scores for normal-hearing listeners in the 6-channel condition is, perhaps, the strongest evidence that some patients are able to extract as much information from six channels as are normal-hearing subjects.

The information extracted from the signals in the present experiment was probably spectral as opposed to temporal information. The signals in the synthetic vowel study were created with equal vowel duration. Temporal factors did not play a role in the recognition of these signals. The multitalker vowels differed normally in duration, but different speakers employed different durations for short and long vowels. The similar level of identification accuracy for the multitalker vowels and synthetic vowels suggests that vowel duration was not a powerful cue to vowel identity in the multitalker-vowel task. The identification of consonant place of articulation was, more likely, also based on spectral information because there are few strong temporal cues to place. Overall, the results suggest that some implant patients in the present experiment were able to extract sufficient spectral information to allow them to identify speech signals within the range of identification accuracy established by normal-hearing subjects listening to speech processed through six channels.†

Not all patients achieved scores within the range of normal. One factor that characterizes the patients with relatively poor identification scores is a relatively small dynamic range (10 to 14 dB averaged over electrodes). In contrast, three out of the four patients with the highest identification scores in the test the least confounded by practice, i.e., multitalker vowels, have a dynamic range of 19 dB or greater averaged over electrodes. One of these subjects had a 12 dB dynamic range. In unpublished experiments with normal-hearing subjects listening to signals processed through 6-channel systems, a 23 point reduction in percent correct scores for multitalker vowels was found when the signals were linearly compressed into a 12 dB range. This outcome is reasonable from the view that the location of spectral peaks and valleys in the input spectrum is coded by relative differences in rms levels across electrodes. A reduction of differences in rms levels among electrodes reduces the distinctiveness of spectral peaks and valleys, which, in turn, reduces the distinctiveness of the spectral cues to vowel and consonant identity. The relatively poor scores for patients with relatively small dynamic ranges and the relatively good scores for patients with relatively large dynamic ranges are consistent with this point of view. However, the outcome that one patient with a small (12 dB) dynamic range achieved a high score on the test with multitalker vowels demonstrates that dynamic range is not the only factor affecting performance.

The patients in this experiment were not a random sample of patients with 6-channel CIS processors, and, thus, the proportion of implant patients falling within the range of performance of normal-hearing listeners in this experiment does not predict that proportion in a larger, random sample. The three patients who achieved the highest scores in the present experiment have achieved scores of greater than 70% correct on the NU-6 word test. The number of patients who achieve this level of performance with any implant is relatively small.

We note that the cochlear implant patients had at least a month, and in a few cases several years, to adapt to 6-channel stimulation. The normal-hearing listeners had only several hours of experience listening to stimulations of processors with two to nine channels. If normal-hearing listeners were allowed months or years of practice with a 6-channel processor, it is possible that scores on tests of speech recognition would increase. This possibility constrains the generality of the findings reported above.

Conclusion TOP

Some cochlear implant patients who use a 6-channel CIS processor and a monopolar electrode array extract sufficient spectral information from speech signals to achieve scores on tests of speech identification that are within the range of scores established by normal-hearing subjects listening to speech processed through six channels. The ability to achieve this level of speech recognition performance may be related to a patient's dynamic range. A larger dynamic range may allow better resolution of spectral detail and, as a consequence, better speech recognition.

Acknowledgments: TOP

This research was supported by grant DC RO1 00654-7 from the National Institute on Deafness and Other Communication Disorders.

References TOP

Dorman, M. F., Dankowski, K., Smith, L., & McCandless, G. (1989). Identification of synthetic vowels by patients using the Symbion multichannel cochlear implant. *Ear and Hearing, 10*, 40-43.
[Context Link]
Dorman, M. F., Loizou, P., & Rainey, D. (1997). Speech recognition as a function of the number of channels of stimulation for processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America 102*, 2403-2411.
[Context Link]
Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K.(1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099-3111.
[Context Link]
Lawson, D., Wilson, B., & Finley, C. (1993). New processing strategies for multichannel cochlear prostheses. In J. H. J. Allum, D. J. Allum-Mecklenburg, F. P. Harris, & R. Probst (Eds.),*Progress in brain research*, 97 (pp. 313-321). New York: Elsevier Science Publishers B.V.
[Context Link]
Miller, G., & Nicely, P. (1955). Analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-353.
[Context Link]
Shannon, R., Zeng, F-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303-304.
[Context Link]
Tyler, R., Gantz, B., Woodworth, G., Parkinson, A, Lowder, M., & Schum, L. (1996). Initial independent results with the Clarion cochlear implant. *Ear and Hearing, 17*, 528-536.
[Context Link]
Tyler, R., Preece, J., & Tye-Murray, N. (1986). *The Iowa phoneme and sentence tests*. Iowa City, IA: The University of Iowa.
[Context Link]
Van Tassel, D., Greenfield, D., Logemann, J., & Nelson, D. (1992). Temporal cues for consonant recognition: Training, talker generalization and use in evaluation of cochlear implants. *Journal of the Acoustical Society of America, 92*, 1247-1257.
[Context Link]
Wilson, B. S., Finley, C., Lawson, D., Wolford, R., Eddington, D., & Rabinowitz, W. (1991). Better speech recognition with cochlear implants. *Nature, 352*, 236-238.
[Medline Link] [CrossRef] [Context Link]

Wilson, B., Lawson, D., & Zerbi, M. (1995). Advances in coding strategies for cochlear implants. *Advances in Otolaryngology-Head and Neck Surgery, 9*, 105-129.
[Context Link]
Zierhofer, C., Peter, O., Bril, S., Pohl, P., Hochmair-Desoyer, I., & Hochmair, E. (1994). A multichannel cochlear implant system for high-rate pulsatile stimulation strategies. In I. Hochmair-Desoyer and E. Hochmair (Eds.), *Advances in cochlear implants* (pp. 204-207). Mantz: International Interscience Seminars.
[Context Link]

Reference Notes TOP

1 Brill, S., Gstottner, W., Helms, J., v. Ilber, C. H., Baumgartner, W., Muller, J., & Kiefer, J. (1996). *Optimization of channel number and stimulation rate for the fast CIS strategy in the COMBI 40+*. Paper presented at the International Workshop on Cochlear Implants, Vienna, Austria, October 24-25.
[Context Link]

*We are indebted to an anonymous reviewer for this suggestion. [Context Link]

†The normal-hearing listeners in Dorman et al. (1997) also were tested with processors that output noise bands similar to those in Shannon et al. (1995). The conclusion that some cochlear implant patients achieve scores within the range of normal holds also when normal-hearing listeners are tested with processors that output bands of noise. [Context Link]