CONTINUING IMPROVEMENTS IN SPEECH PROCESSING FOR ADULT COCHLEAR IMPLANT PATIENTS

R. D. Hollow, BSc, DAUD; R. C. Dowell, PhD, DAUD; R. S. C. Cowan, PhD, DAUD; M. C. Skok, BSc, DAUD; B. C. Pyman, MBBS, FRACS; G. M. Clark, PhD, FRACS

From the Cochlear Implant Clinic, the Royal Victorian Eye and Ear Hospital and the University of Melbourne (Hollow, Dowell, Cowan, Pyman, Clark), and the Cooperative Research Centre for Cochlear Implant, Speech and Hearing Research (Cowan, Skok, Clark), Melbourne, Australia.

INTRODUCTION

The Cochlear 22-channel cochlear implant has employed a succession of improved speech-processing strategies since its first use in an adult patient in Melbourne in 1982. The first patients received the F0F2 coding strategy developed by the University of Melbourne, in the Wearable Speech Processor (WSP). The F0F2 coding scheme presented the implant user with three acoustic features of speech. These were 1) the

after implantation, including etiologic factors such as duration of deafness as percentage of life span and biophysical indicators of poor neuronal survival (Kessler et al, this suppl, section 14). The CID curves (Fig 1) suggested that objective criteria might be used to distinguish discrete patterns of learning in Clarion users. These learning patterns could, perhaps, be recognized early in the course of postoperative treatment in order to identify those patients who would or would not be likely to benefit from intensive training.

In order to categorize these learning curves objectively, we developed an improvement rate factor (IRF) based on the ratio of the improvement of CID score at 3 months versus preoperatively to that eventually achieved at 12 months:

\[
IRF = \frac{CID(3 \text{ mo}) - CID(\text{preop})}{CID(12 \text{ mo}) - CID(\text{preop})}
\]

The distributions of IRFs are shown separately in Fig 3 for patients who ultimately achieved poor, moderate, and excellent CID results. The majority of patients with either excellent or poor results achieved most of their ultimate performance almost immediately after fitting. A small group of patients with moderate to excellent results improved substantially from initially poor results. Of interest, this group was composed exclusively of patients with no prognostic indicators for poor results.

DISCUSSION

The high IRF (80% to 100% in 19 of the 32 patients obtaining moderate to excellent CID results) suggests that the Clarion provides speech percepts that are recognized and interpreted correctly in most of those patients who have reasonably intact auditory pathways. This agrees with subjective impressions that most patients already had substantial open-set speech recognition at their initial fitting, although the first objective measures were not obtained until 3 months. For the more demanding NU-6 test (Fig 2), improvement was more gradual, suggesting that systematic aural exercises might enhance learning. Even for this more difficult test, however, most of the improvement occurred in the first 6 months, with scores leveling off in the latter 6 months, particularly for the better performers. Of the patients with relatively poor CID results at 3 months, a subgroup of patients with poor prognostic factors was readily identified. These patients never achieved substantial NU-6 scores during the 12-month period of this study, although some seemed to be improving very gradually in their CID scores. Longer follow-up is needed for this group. The remaining poor performers at 3 months included an apparently indistinguishable admixture of patients who did not improve substantially, along with those who learned over time to make moderate and occasionally even excellent use of their prosthesis. This would seem to be the most appropriate group to receive intensive aural rehabilitation. It remains to be seen whether improvements in rehabilitative procedures and/or more appropriate speech processor strategies can accelerate the rate at which these patients derive functional benefit from their cochlear prostheses.

In postlingually deaf adult subjects, the IRF presumably relates to the ability of the cochlear prosthesis to reproduce key features of the temporospatial pattern of auditory nerve activity that the subject learned to interpret before becoming deaf. If these patterns are unfamiliar, then the subject must learn what amounts to a new language. Together with the bimodal distribution of speech perceptual results reported elsewhere for the Clarion (Kessler et al, this suppl, section 14), the high IRF suggests that the limiting factor in speech performance in many of these patients is the condition of their peripheral auditory nervous systems, rather than their ability to apply high-level cognition to overcome technological limitations of the device.

REFERENCES

amplitude of the waveform, presented as the amount of current charge, 2) fundamental frequency (F0) or voice pitch, presented as rate of biphasic pulsatile stimulation, and 3) the spectral range of the second formant frequency (F2), which was represented by varying the site of stimulation along the electrode array.

In 1985, a new speech processor, the WSPIII, was introduced, along with the FOF1F2 speech-coding strategy. In this scheme, a second stimulating electrode pair was added to present the spectral range of the first formant frequency (F1). This was followed in 1989 by the Miniature Speech Processor (MSP), which employed the Mpeak coding scheme. This scheme incorporated all of the features of the FOF1F2 scheme, plus information from three high-frequency band-pass filters. Patients using the Mpeak strategy received frequency information from four stimulating electrode pairs with each glottal pulse.

Cochlear Pty Limited have recently developed a new speech-processing strategy, Speak, used in a new-generation Spectra 22 speech processor. This strategy is based on research conducted at the University of Melbourne with the Spectral Maxima Speech Processor (SMP).\(^2\) In the Speak strategy, speech input is passed through 20 programmable band-pass filters. The filter bank is repetitively scanned at an adaptive rate, and the largest outputs of these filters are presented to stimulating electrode pairs. On average, 6 stimulating electrode pairs are chosen per period, with the range being between 1 and 10.

The Melbourne Cochlear Implant Clinic now has over 120 adult patients who have used the various models of Nucleus 22-channel cochlear implant. This paper examines the speech perception performance of the postlingually deaf patients in this group using electrical stimulation alone, with respect to the speech-processing strategy used for the first 3 to 6 months after implantation. Included in this overview are the preliminary results for patients using the new Speak coding strategy from the time their implant was first started up.

**METHOD**

Patients in the Melbourne Clinic routinely undergo speech perception testing postoperatively to

1. Establish that the speech processor is mapped to output the current levels that provide optimum hearing sensations.
2. Ensure that the cochlear implant system is functioning properly.
3. Establish postoperative versus preoperative speech perception improvement.
4. Measure progress in a patient’s performance over time.

To present speech perception scores obtained with the initial speech-processing strategy used by each patient, results were taken from tests administered at 3 to 6 months after implant “switch-on.” For each patient, we obtained: phoneme scores on an open-set word test such as Consonant-Nucleus-Consonant (CNC)\(^3\) or Arthur Boothroyd (AB)\(^4\) words, and open-set sentence scores using Central Institute for the Deaf (CID)\(^5\) everyday sentences. The tests were presented free-field by means of recorded materials, with the patients using their cochlear implant alone, without the assistance of lip-reading.

It should be noted that Mpeak data contain scores from both CNC and AB word lists. The AB word data were included to maximize the sample size, as only 16 of the 29 patients tested with monosyllabic word materials had scores on the CNC word test. The authors appreciate that although both tests can be scored by the number of phonemes correctly recognized, the word lists do have their differences, and this will contribute to score variability.

The patients included in this study were all 17 years or older at the time of testing, were English-speaking, had a postlingual onset of severe to profound deafness, and had an implant operating with 12 or more active electrodes.

**RESULTS**

For each group of patients tested with a particular speech-processing strategy, the mean open-set sentence and speech phoneme scores were calculated. The Figure represents the mean electrical stimulation alone scores for both CID Everyday sentences and CNC and AB word tests.

The mean sentence scores were 15.9% with the FOF2 strategy (N = 13), 38.5% with FOF1F2 (N = 32), and 59.1% with Mpeak (N = 27), and the preliminary result with the Speak strategy (N = 6) was 82%. An independent samples t test (with unequal variances) indicated the FOF2 to FOF1F2 to Mpeak improvements to be statistically significant (p < .01). The Mpeak to Speak improvement was also significant (p < .05). Within the results for each speech-processing group, there were wide ranges of scores: 0% to 58% with FOF2, 2% to 98% with FOF1F2, 1% to 100% with Mpeak, and 33% to 100% with Speak.

As with the sentence scores, the mean phoneme scores from the word tests increased significantly (p < .01 for FOF2 to FOF1F2 and FOF1F2 to Mpeak, and p < .05 for Mpeak to Speak) with each improved coding strategy: 23.1% with FOF2 (N = 13), 33.2% with FOF1F2 (N = 28), 47.9% with Mpeak (N = 29), and 69.3% with Speak (N = 6). Again, there was a wide range of scores for each strategy: 6% to 39% with FOF2, 12% to 69% with FOF1F2, 3% to 85% with Mpeak, and 38% to 93% with Speak.

**DISCUSSION**

To provide an overview of the open-set speech perception...
performance of adult patients implanted at the Melbourne Cochlear Implant Clinic since 1982, it is necessary to examine the influence that the use of different speech-processing strategies has had on mean speech perception performance in a group of postlingually deaf adults.

The results show that for each speech-processing strategy introduced, there has been an increase in the mean scores obtained on both open-set word and sentence tests. This suggests that the additional spectral and temporal cues provided by including first formant amplitude and frequency information in the case of the FOF2 to FOF1F2 upgrade, and the amplitude and frequency information in three high-frequency band-pass filters in the case of the FOF1F2 to Mpeak upgrade, have, on average, provided people with an improved representation of speech. This confirms results of earlier studies that compared FOF2 versus FOF1F2 and FOF1F2 versus Mpeak in adults changing from one strategy to another.

The data available from patients using the Speak coding strategy, although limited at present, are also encouraging in this sense. They at least show that there is potential for patients to understand a substantial amount of connected speech using the cochlear implant alone, without the assistance of lipreading, and to receive an improved representation of speech phonemes with this more sophisticated speech-coding strategy. Further studies are necessary, though, to evaluate the Speak strategy on a larger number of cochlear implant users.

Another point of interest is the wide range of speech perception performance present with all strategies. Many factors may influence the final outcome of cochlear implantation. Experience has shown that low speech perception performance can be correlated with a long period of profound deafness, extensive neural degeneration (as estimated with the electrical stimulation of the promontory test), minimal residual hearing prior to implantation, and a reduced number of active electrodes in a patient's map. Whatever the reasons for the variation in speech perception performance, it does indicate that even with more sophisticated speech-processing strategies, there remain some patients who do not obtain substantial open-set speech recognition. With increased experience using these strategies and improved ability to predict the outcome of cochlear implantation, we should aim to provide appropriate preoperative counseling on the anticipated outcomes for patients.

REFERENCES


TONE PERCEPTION OF MANDARIN-SPEAKING POSTLINGUALLY DEAF IMPLANTEES USING THE NUCLEUS 22-CHANNEL COCHLEAR MINI SYSTEM

T.-S. HUANG, MD; N.-M. WANG, MS; S.-Y. LIU, BA

From the Department of Otolaryngology, Chang Gung Medical College, Taipei, Taiwan, Republic of China.

Because Mandarin Chinese is a tonal language, testing the patient's ability to distinguish among four tones is of paramount importance. This paper evaluates the efficacy of the Nucleus 22-Channel Mini System for Mandarin Chinese by comparing the postoperative tone perception test results with the results of the closed-set monosyllable, trochee, and spondee (MTS) test, and the open-set phonetically balanced (PB) word and sentence comprehension tests, which also incorporate tonal features, for eight Mandarin-speaking postlingually deaf patients implanted with this device. Except for one patient, the data clearly indicate that patients who had increased scores on tone perception after implantation also had improvements on other test batteries. Results also substantiate our previous observation that the Nucleus 22-channel cochlear implant enables profoundly and totally deaf patients to distinguish four separate tones in Mandarin Chinese. This would seem to support our earlier speculation that the acoustic cues of fundamental frequency of the four Mandarin tones are extracted by the Multipeak coding strategy of the 22-channel device and transferred to the cochlea, where they are perceived as rate pitch.

INTRODUCTION

Since Mandarin Chinese relies on four separate tones to distinguish words, the success of a cochlear prosthesis depends on the ability of the device to transmit these characteristics. Cochlear implants designed for Western languages did not incorporate this capability in their engineering; therefore, the question of their efficacy and adaptability for Mandarin-speaking patients arises.

In an earlier report, we concluded that although the Multipeak coding strategy of the Nucleus 22-Channel Mini System could benefit the speech and auditory performances of profoundly to totally deaf adult Mandarin-speaking patients, the overall tone perception results were not as favorable as the scores for other closed- or open-set test batteries. In the present study, we further investigate the efficacy of this Multipeak coding strategy with regard to tone perception by comparing
Author/s: Hollow, R. D.; Dowell, R. C.; Cowan, R. S. C.; Skok, M. C.; Pyman, B. C.; Clark, Graeme M.

Title: Continuing improvements in speech processing for adult cochlear implant patients

Date: 1995


Persistent Link: http://hdl.handle.net/11343/27469