Advances in Cochlear Implants
edited by I. J. Hochmair-Desoyer and E. S. Hochmair
Manz, Wien © 1994

Mechanisms of Formant Frequency Resolution with the Ineraid Cochlear Implant

Michael F. Dorman
Speech and Hearing Science, Arizona State University
Tempe, Arizona 85287-0102 USA

Luther Smith
Sinai Samaritan Medical Center
Milwaukee, Wisconsin 53233 USA

James L. Parkin
University of Utah School of Medicine
Salt Lake City, Utah 84132 USA

Abstract **In experiment 1, vowels were created with conflicting temporal and place cues to vowel identity. The results of an identification task indicated that patients relied on information contained in the distribution of energy, or place, cues, for vowel identification. In experiment 2, the discrimination of frequencies in the domain of the first formant was assessed in a three-tone task in which the patients identified the high and low members of a triad. Performance on the test of frequency discrimination was correlated significantly (.70 - .82) with performance on tests of speech understanding. In experiment 3, frequency discrimination was assessed through the Ineraid analogue processor and through a 6 channel, CIS processor. Frequency discrimination was enhanced with the CIS processor.**

EXPERIMENT 1. The evidence from a number of studies on vowel recognition suggests that patients who use the Ineraid cochlear implant obtain a relatively large amount of information from the first formant and much less information from the second formant (Dorman et al., 1989; Eddington, 1983). The research described here was directed to the question of how information about the first formant is coded.

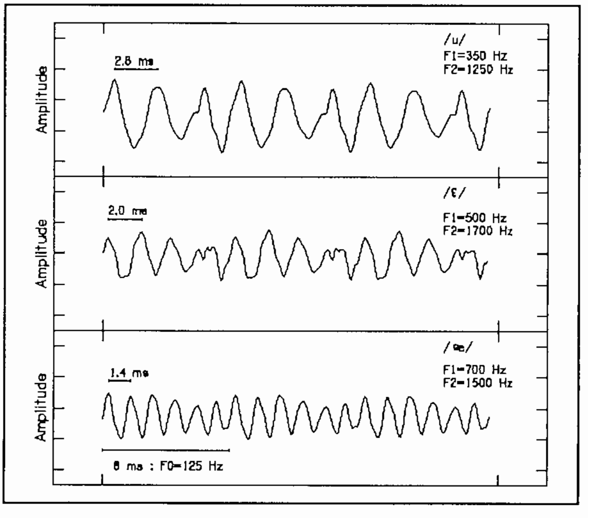

Figure 1. Waveforms from channel 1 of the Ineraid for the vowels /u/, /ε/ and /æ/.

One possibility is that first formant information is coded by rate pitch or temporal cues. This possibility receives some support from the performance of patients fitted with the Vienna single-channel prosthesis on tests of vowel recognition. Vienna patients, like Ineraid patients, rely on information from the first formant for vowel identification (e.g., Rosen and Ball, 1986; Tyler, Tye-Murray, Moore and McCabe, 1989; Wallenberg, Hochmair-Desoyer and Hochmair, 1990). Given the presence of only a single channel, information about the first formant must be derived from temporal cues. Thus, temporal information can be a sufficient cue for recognition of the first formant. Since temporal information about the first formant is preserved in the output of the Ineraid processor (see Figure 1), Ineraid patients may be able to use this information for vowel recognition.

On the other hand, Ineraid patients receive some information from the second formant and it is very unlikely that this information could be coded in terms of temporal cues. Rather, the distribution of energy across electrodes 2-4 must provide a place cue to frequency. Parsimony might dictate that we invoke the same mechanism for coding the first and second formants within the same device. Thus, we could argue that both F1 and F2 are coded by the distribution of energy across the electrode array -- that is, a place code.

We have conducted an experiment to provide evidence on this issue (Dorman et al., 1992). We first modified signal presentation through the Ineraid by adding a stage of gain and attenuation following the Ineraid processor. Thus, signals were input to the Ineraid via the auxiliary input jack, filtered by the Ineraid and then were directed to the digital gain and attenuator stage. Finally, the four channels of information were directed to voltage to current converters and optical isolators before being delivered to the patients electrodes via the pedestal. This arrangement allowed us the modify the amplitude of a channel without altering the temporal information within that channel.

The nature of the temporal cues, shown as the output of channel 1, for the three vowels /u ɛ æ/ used in the experiment are shown in Figure 1. The frequency of the first formant -- 350 for /u/, 500 for /ɛ/ and 700 for /æ/ -- is clearly carried by the temporal waveform. That is, the intervals between these waveform components change from 2.8 ms to 2.0 ms to 1.4 ms.

The nature of the place cues -- in terms of the signal levels at each electrode for each of the vowels -- are shown in Figure 2. The pattern of energy for each vowel is different.

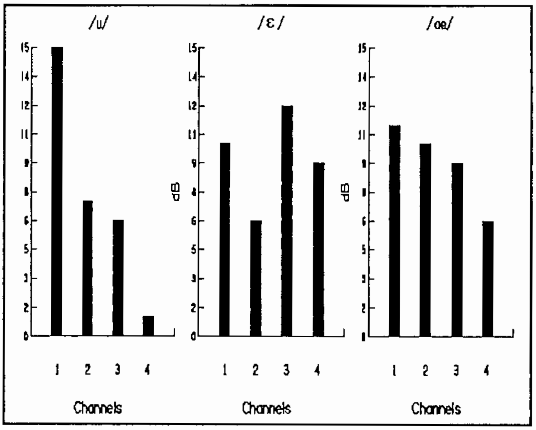

Figure 2. Relative signal levels in channels 1 - 4 for the vowels /u/, /ɛ/ and /æ/.

In our experimental manipulation, we took the signals from channels 1 - 4 from, for example, /u/, and with the aid of the gain and attenuator stage, altered the signals to have the amplitude characteristics appropriate for one of the other two vowels. To give /u/ the channel amplitude characteristics of /ɛ/, we reduced the amplitude of channel 1, reduced 2 only a little and increased 3 and 4 a large amount. Manipulations of this kind were carried out for all three vowels.

Thus, we created signals with conflicting sources of information about formant frequency. One source was the temporal information in each channel. The other was the place information carried by the distribution of energy across the electrodes. At issue was whether patients would base identification on the temporal or the place cues.

The results for the six subjects are shown in Table 1. The stimuli which were not altered were identified with high accuracy -- over 90 % correct. When we altered the pattern of energy across the electrodes, the identification of each stimulus was altered significantly. The altered stimuli were identified, most generally, on the basis of the pattern of energy across the electrode array. For example, when we gave the /ɛ/ stimulus the amplitude characteristics of /u/, patients indicated that they heard /u/ 93% of the time. When we gave /ɛ/ the amplitude characteristics of /æ/, then /æ/ was heard the majority of the time (67%). This pattern, to a greater or lesser degree, is found in the rest of the data. Most generally, patients based vowel identification on the distribution of energy across the electrode array and not on the temporal information in the channels. Thus, it appears that Ineraid patients extract information about the frequency of the first formant of speech signals by the distribution of energy across electrodes 1 and 2. They do not appear to use rate pitch or temporal encoding for the first formant. Thus, if patients have but a single channel of information, as with the Vienna device, then they can use temporal information to code the first formant. However, when cues to cochlear place of stimulation are also available, as with the Ineraid device, then temporal cues are not used for identification of the frequency of the first formant.

Table I

	Response		
Stimulus	"u"	"ɛ"	"æ"
/u/	93	3	3
/ɛ/	0	97	3
/æ/	0	3	97
/u/ → /ɛ/	0	84	16
/u/ → /æ/	0	63	37
/ɛ/ → /u/	93	3	3
/ɛ/ → /æ/	10	23	67
/æ/ → /u/	90	7	3
/æ/ → /ɛ/	0	47	53

EXPERIMENT 2. In the course of our studies on the mechanism of coding the first formant, we have conducted studies of frequency discrimination in the region 250 Hz to 950 Hz. At issue was the relationship between frequency discrimination for signals in this frequency domain and speech understanding. To assess frequency discrimination, patients were

presented, through their signal processors, three signals in a sequence, for example, 250, 350 and 450 Hz, and were asked to indicate which was the highest frequency and which was the lowest frequency. The order of the stimuli within a triad was randomized across trials and loudness was balanced among members of each triad before testing. Each triad was presented five times for identification. As shown in Figure 3, some patients (patient 1 : filled circles) were able to resolve the 100 Hz differences very well -- 90 % or better accuracy out to the triad centered at 750 Hz. Other patients (patient 14: open circles), unfortunately, did not perform as well. To assess the relationship between this measure of low frequency resolution and speech understanding, we computed the correlations between performance on the frequency discrimination task (in terms of percent correct) and scores on several tests of speech identification. For 25 patients, the correlation between frequency discrimination and vowel identification was .70. For spondee recognition it was .74; for single syllable (NU-6) word identification it was .81; and for consonant identification it was .82. Thus, we find that low frequency discrimination accounts for half, or slightly over half, of the variance in vowel, consonant and word identification for patients who use the Ineraid cochlear implant. It remains to be seen whether similar tests at higher frequencies, in the domain of the second formant, will produce similar results.

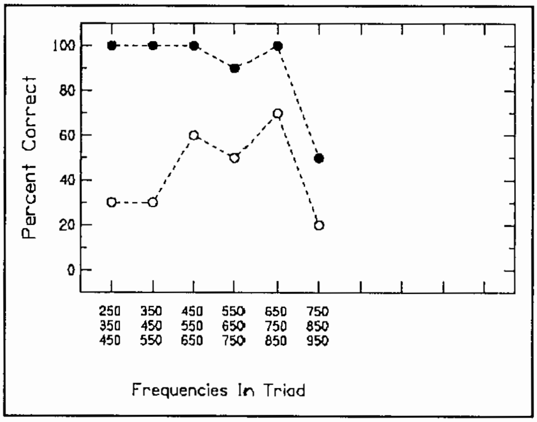

Figure 3. Percent correct identification of frequencies in a triad as a function of the center frequency of the triad. The performance of two patients is illustrated.

EXPERIMENT 3. It is to be expected that CIS processors (Wilson et al., 1991) will replace the analogue signal processors currently used by Ineraid patients. Wilson et al. report significantly higher scores for Ineraid patients on tests of speech understanding when using a CIS processor than when using an analogue processor. We should expect that one factor, perhaps the principle factor, underlying the improved performance is improved frequency discrimination. To assess whether this is the case, triads were constructed with varying frequency differences between the members of the triad. In Figure 4, percent correct identification of a triad is plotted as a function of the differences in frequency among members of the triad. The center frequencies varied from 400 Hz to 2.0 kHz. The independent variable was processor type: Ineraid vs a 6-channel CIS processor. The patient was one of the best Ineraid patients in terms of speech understanding in the United States. He has participated in many psychophysical experiments. His

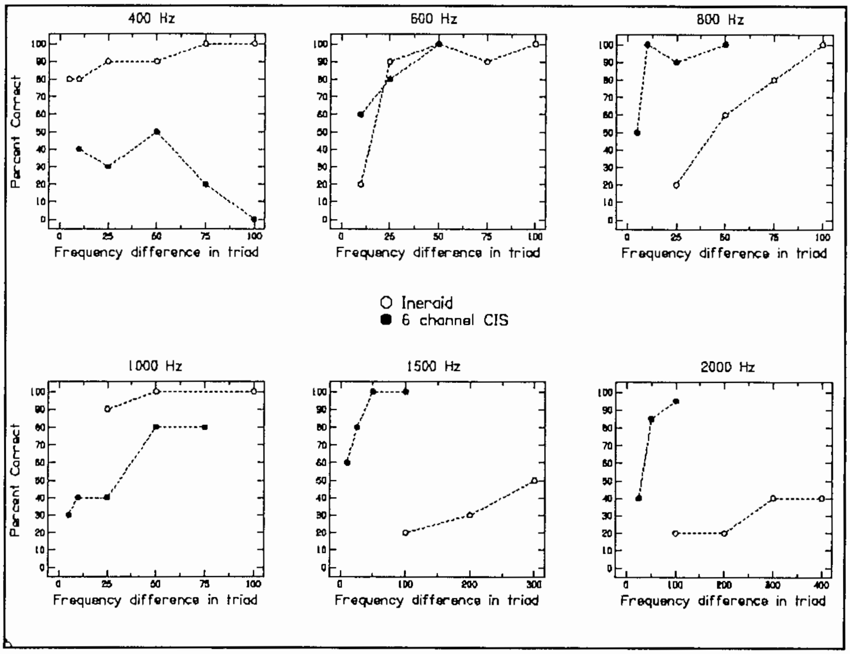

Figure 4. Frequency discrimination as a function of triad center frequency for a patient tested with the Ineraid electrode array linked to the Ineraid, analogue processor and to a 6-channel CIS processor.

performance on the tests of frequency discrimination may not be typical of Ineraid patients with less experience in psychophysical testing. None the less, we suspect that this patient's performance illustrates the differences in frequency discrimination allowed by the different signal processing strategies.

Frequency discrimination with the Ineraid is very good for frequencies in the domain of the first formant. At 400 Hz, the point of 75% accuracy was reached with less than a 5 Hz difference among members of the triad. As the center frequency increased, the point of 75% accuracy increased -- at 600 Hz it was about 23 ms; at 800 Hz it was about 58 Hz; at 1 kHz it was less than 25 Hz; at 1500 and 2000 Hz it was greater than 300 Hz. When signals were presented through the CIS processor, the point of 75% accuracy at 2 kHz was about 45 Hz; at 1500 Hz it was about 23 Hz; at 1 kHz it was about 48 Hz; at 800 Hz it was about 8 Hz; at 600 Hz it was about 22 Hz. Only at 400 Hz was discrimination poor with the CIS processor (performance in this region is the subject of ongoing research). Thus, for the set of frequencies between 600 Hz and 2 kHz used in this experiment, 75% accuracy in discrimination was reached between 8 and 40 Hz. This level of frequency discrimination is, perhaps, the principal reason why this patient achieves scores of 90 % or better NU-6 identification with a 6-channel CIS processor.

REFERENCES

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

M. Dormam, L. Smith, M. Smith & J. Parkin. The coding of vowel identity by patients who use the Ineraid cochlear implant. Journal of the Acosutical Society of America, 92, 3428-3431, 1992.

D. Eddington. Speech recognition in deaf subjects with multichannel intracochlear electrodes. In C. Parkins and S. Anderson (Eds.), Cochlear prostheses: An international symposium. Annals of the New York Academy of Sciences, 405, 241-258.

S. Rosen & V. Ball. Speech perception with the Vienna extra-cochlear single-channel implant: a comparison of two approaches to speech coding. British Journal of Audiology, 20, 61-83, 1986.

R. Tyler, N. Tye-Murray, B. Moore, & B. McCabe, B. Synthetic two-formant vowel perception by some of the better cochlear-implant patients. Audiology, 28, 301-315, 1989.

B. Wilson, D. Lawson, C. Finley, D. Eddington & W. Rabinowitz. Better speech recognition with cochlear implants. Nature, 352, 236-238, 1991.

ACKNOWLEDGEMENT

This research was supported by grant DC-000654 from the National Institute of Deafness and Other Communicative Disorders. Experiment 4 was conducted with the assistance of B. Wilson, D. Lawson, C. Finley and M. Zerbi. We thank MP for many hours of testing without complaint.