Speech Understanding with Dichotic Presentation of Channels: Results from Acoustic Models of Bilateral Cochlear Implants

Michael F. Dorman[1], Philip C. Loizou[2], Anthony J. Spahr[1], Erin S. Maloff[1] and Susan Van Wie[1]

[1]Arizona State University : [2]University of Texas at Dallas.

Background

During the past year we have tested two patients fit with binaural implants -- an Ineraid electrode array / Med El CIS-Link processor for one ear and a Clarion Hi-Focus system for the other. Both patients showed a diotic advantage for speech in noise presented through a single, center facing loudspeaker and, thus, appeared to be able to integrate information from the two ears. This result prompted a series of experiments with acoustic models of bilateral implants. In our prototypical experiment even numbered channels of an *N*-channel processor were presented to one ear and the odd numbered channel were presented to the other. The aim of these experiments was to probe the differences in intelligibility between monaural and dichotic presentation of channels from a speech processor. In a second experiment we shifted the output frequencies of one set of channels to assess the effect of having one set of channels at the 'wrong' frequencies. At issue was how how far 'off' in frequency a dichotically-presented channel could be and still contribute to diotic summation.

Experiment 1

Method

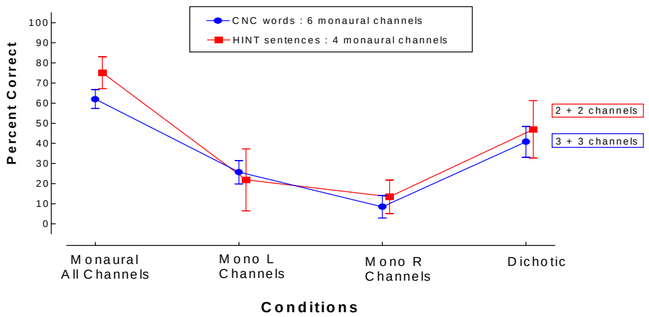
Subjects: The subjects were normal-hearing undergraduate students.

Stimuli: CNCs and HINT sentences

Signal processing: To create signals with similar levels of monaural intellibility, the CNC words were processed into 6 channels and the HINT sentences into 4 channels. To create these stimuli, the signals were first processed through a preemphasis filter (high-pass filter with a 1200 Hz cutoff frequency and 6 dB/octave rolloff) and then bandpassed into *N* frequency bands (where *N* varied from 4-7) using sixth-order Butterworth filters. The spacing of the frequency bands was logarithmic. The envelope of the signal was extracted by full-wave rectification and low-pass filtering (second-order Butterworth) with a 400 Hz cutoff frequency. Sinusoids were generated with amplitudes equal to the root-mean-square (rms) energy of the envelopes (computed every 4 ms) and frequencies equal to the center frequencies of the bandpass filters. The sinusoids were summed and presented to the listeners at a comfortable level.

Four stimulus conditions were created : monaural presentation of signals, monaural partial channels to the left and right ears, and dichotic presentation of signals. For the partial channel conditions the odd numbered channels were presented to the left ear and the even numbered channels to the right ear. For the dichotic conditions the left and right, partial-channel stimuli were output simultaneously.

Results



Outcome

The principal finding of Exp. 1 was that for CNCs and HINT sentences dichotic presentation of channels led to significantly poorer performance than monaural presentation of channels. For this type of stimulus material, subjects differed greatly in their ability to combine information from the two ears. While some subjects achieved scores in the dichotic conditions that were nearly the equal of scores in the monaural condition, others achieved scores only slightly better than monaural, partial-channel scores. If similar individual differences exist in patients fit with bilateral cochlear implants, then this factor, independent of pathophysiology, may account for some of the variance in the scores acheived with bilateral processors.

Experiment 2

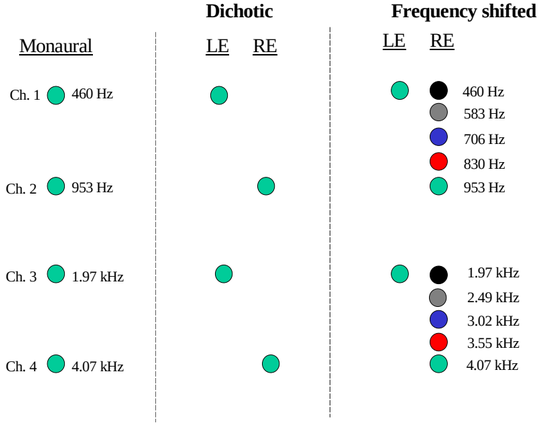
The aim of Exp. 2 was to determine how far 'off' in frequency a dichotically presented channel could be and still contribute to diotic summation.

Subjects: The subjects were normal-hearing undergraduate students.

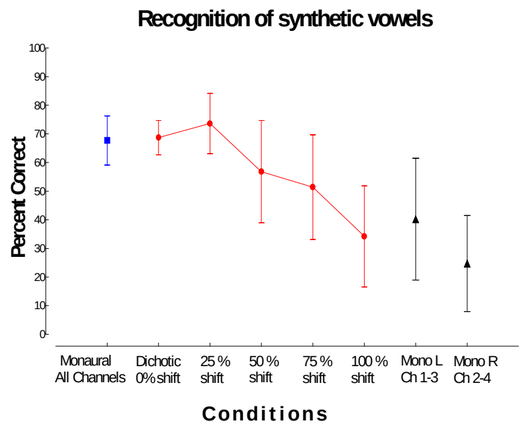
Stimuli: Synthetic vowels in 'bVt' format and HINT sentences.

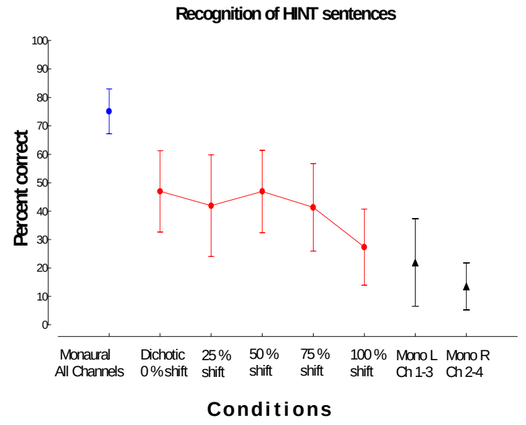
Signal processing: Signals were processed in the manner described in Exp. 1. To create frequency shifted signals, the even numbered channels, presented to the right ear, were shifted in frequency 25%, 50%, 75% and 100 % of the distance between the 'correct' output frequency and the output frequency of the next lower channel. For example, the appropriate output frequency for channel 2 (for 4-channel HINT sentences) was 953 Hz. In the 25, 50, 75 and 100 % shift conditions the output frequencies were 829,706, 583 and 460 Hz, respectively. ***In the 100 % shift condition, the output frequencies of channels 2 and 4, presented to the right ear, were identical to the output frequencies of channels 1 and 3, presented to the left ear. The signals differed only in time waveforms.***

Stimulus Conditions



Results





Outcomes

For vowels: Relative to the dichotic condition the 75% and 100 % frequency shift conditions produced significantly poorer scores. The 50 % shift produced a decrease of marginal significance. Relative to the best partial-channel, monaural condition only the dichotic and 25 % shift conditions produced significantly better scores.

For sentences : Relative to the dichotic condition only the 100 % frequency shift condition produced significantly poorer scores. Thus, there was no effect of up to 75% shift in channel location. The 100 % shift condition did not differ from the best partial channel condition.

Conclusions

We used vowels and sentences as stimulus materials in Experiment 2 because they offered a contrast in the use of frequency and amplitude-envelope information. Frequency based information is necessary for the recognition of both vowels and words in sentences. In contrast, envelope information was of little or no use in identifying the synthetic vowels used in this experiment but was of every use in identifying words in the HINT sentences. Thus, we expected a different outcome for the vowel and sentence materials. The expected was obtained: For vowels, frequency information needed to be within 25 % of where it 'should have been' in order to allow normal summation of information across ears. For sentences, even a 75 % shift in frequency did *not* produce a decrement in performance. Thus, getting channel frequencies 'correct' in bilateral stimulation is very important for vowels but is less so for sentences. For sentences, useful envelope information can be obtained even from signals delivered far from the appropriate cochlear location. This outcome may account for some of the success of bilateral implants.

Summary

- Dichotic presentation of channels commonly leads to poorer speech recognition than monaural presentation of the same number of channels.
- Individuals differ greatly in the ability to 'fuse' channel information presented to the two ears. This may account for some of the variance in the success of bilateral implants.
- Presenting dichotic channels at, or at least near, the 'correct' place of stimulation is important for vowels but not sentences.
- Amplitude information can be obtained from channels presented at frequency locations far from where they 'should be.' This factor may account for some of the success of bilateral implants. That is to say, while some information is lost from poorly positioned channels in the frequency domain, envelope information even from poorly positioned channels can be of use in speech recognition.