T. G. Thomas, Experimental evaluation of improvement in speech perception with consonantal intensity and duration modification, Ph. D. Thesis, Department of Electrical Engineering, Indian Institute of Technology Bombay, 1996.

Supervisors: Prof. P. C. Pandey and Prof. S. D. Agashe

Abstract - A promising scheme for enhancing the speech signal is based on studies of speaking clearly for the hearing impaired. Studies of the differences between "clear" speech and "conversational" speech have identified certain consistent acoustic modifications of the speech signal in clear speech. Two characteristics of clear (intelligible) speech, namely consonant-to-vowel intensity ratio (CVR) and consonant duration (CD), have been chosen for evaluation in the present investigation.

The aim of the experiments reported here was to study the extent of CVR modification and CD modification that would be helpful in speech perception by the hearing impaired. A computerized test administration system was developed in order to automate the experiments. It was decided to use synthetic stimuli since the segmentation of individual phonemes would be accurate, and it would be possible to manipulate the various acoustic segments independently. The stimuli involved stop consonants in the consonant-vowel (CV) and vowel-consonant (VC) contexts of various vowels. The subjects were normal-hearing persons and hearing impairment was stimulated by mixing each stimulus with masking broadband noise. The results each experimental run were obtained in the form of a stimulus-response confusion matrix. The confusion matrices were evaluated on the basis of recognition scores and information transmission analysis. Average response time was also considered as another possible measure of comparing the test stimuli processed differently.

The results show that increasing CVR does improve recognition scores. Some vowel confusions were observed in the VC context at higher CVRs, but as long as CVR modification was restricted to about 9-10 dB there was no adverse effect on the recognition of vowels. The information transmitted on the basis of feature-classification, as well as the average response times were also found to improve with increasing CVR.

For the CD modification experiments, the acoustic segments that constitute the consonant phonemes were altered in duration separately to study their individual effects. The results suggested that at higher noise levels, a formant transition duration modification of upto about 50% may be combined with burst duration modification and expected to yield better performance. Voice onset time (VOT) does not appear to be a suitable parameter for modification as performance decreased with increasing VOT.

These results suggest that it would be fruitful to develop a scheme that can identify the boundaries of various phonetic segments and subsegments and then perform appropriate modifications in C/V ratio, formant transition duration, and burst duration. This scheme could then be tested on the hearing impaired.