

Speech Recognition : Study of Fundamental Frequency as a Variable

Ashok Kumar

Student, A.I.I.S.H., Mysore.

Computer has invaded all walks of life of human beings. Speech Pathology is no exception to this. Computer has been used extensively for Speech analysis and synthesis for the diagnosis and treatment of speech disorders. However, speech recognition has not been used well for the purpose of speech and language therapy, even though it has lot of potentials in terms of providing reinforcement to the cases, motivation and drill to the cases. Presently some of the attempts have been made to use this for articulation testing and therapy. In this context it was necessary to review the literature regarding speech recognition and also study possible variables affecting speech recognition. Therefore, this is part of a proposed extensive study on application of speech recognition in speech therapy and diagnosis.

The present study was aimed at reviewing the relevant literature and answering the following question. :-

1. Is there any difference in terms of recognition of digits when the pitch is varied from the habitual pitch?
2. Is there any difference in terms of three methods of speech recognition (Cosh Measure, MPR & LPC i.e., Euclidean Distance Measure) in terms of recognition of digits?

It was decided to use digits to start with, to make the problem simpler and also as others had used digits for recognition.

A sample comprising of five subjects (all males) with no history of hearing loss and vocal pathology were taken for this study.

The experiment was conducted in two phases. In the first phase the subjects were asked to utter the digits (0 to 9) using their habitual pitch at normal loudness. In the second phase of the experiment the subjects were asked to utter the same digits (0 to 9), in three different pitches i.e., habitual, high and low pitch. The recorded speech samples were analyzed with the help of a computer (PC-XT). For recognition of digits the methods, Cosh Measure, Minimum Prediction Residual and Linear Prediction Coefficients were used.

The presented study has revealed that the percentage of correctly recognized digits is varied with respect to habitual pitch. The digits uttered at low pitch were very poorly recognized (about 40%). Some what better result was seen at higher pitch level with respect to low pitch level. But when compared to habitual pitch the digits were less recognized at high pitch.

A comparison was also made between the three methods of speech recognition used in this study (Cosh Measure, MPR and Linear Prediction Coefficient). On comparison, it was found that there was difference between the three methods of recognition. But the methods Cosh Measure and Linear Prediction Coefficient (Euclidean Distance Measure), yielded almost same result; whereas the Minimum Prediction Residual method yielded very poor result.

Above all it was found that Cosh Measure is reliable method of speech recognition. Similar results have been reported by Gray and Markel (1976).

Based on the above results it may be inferred that the variation in pitch affects the recognition of digits.

This finding however, is restricted to small sample. Further studies need to be carried out before generalizing these results.

Recommendations for future research :

1. Similar experiments can be carried out with different speech samples.
2. This study can be done on larger population.
3. Other recognition programmes may be used with similar stimuli and conditions.