Automatic Speech Recognition Performance of Alveolar Plosive and Retroflex Plosive Phonemes of Malayalam Language

Dr. Cini Kurian

Department of Computer Science , Al-Ameen College, Edathala , Aluva , Kerala

Corresponding Author: Dr. Cini Kurian

Abstract: Speech recognition applications are becoming common and useful in this day and age as many of the modem devises are designed and produced user-friendly for the convenience of general public. The main objective of this paper is to explore the speech recognition performance Alveolar Plosive and Retroflex plosive phonemes of Malayalam language.

I. INTRODUCTION

Speech signal has spectral characteristics which differ dynamically within a stretch of a few milliseconds. These temporarily varying spectral chunks can be grouped into abstract classes called allophones in phonetic study. The correlation of the phonemes and letters , especially , in the Indian languages, make the speech research fairly easy compared to English. But in India, these concept cannot be similar for all languages. Many of the phonemes has similar realization. Phonetic realization[1] of all the phonemes will not be the same across languages. Some language has some peculiar phonemes and some phonemes has different realization.. Malayalam[2], a language that has not been the subject of much research in the speech recognition as well as phonetic and phonological literature. So in this paper , tried to explore some phonemes of Malayalam language which has unique phonetic realization or has unique properties along with its speech recognition performance.

II. METHODOLOGY

Malayalam has 52 consonant phonemes, encompassing 7 places of articulation and 6 manners of articulation, as shown in Table 6.1 below. In terms of manner of articulation, plosives are the most complicated, for they demonstrate a five-way distinction in bilabials, dentals, alveolar-palatals, retroflex, and velars. A bilabial plosive, for example, is either voiceless or voiced. Within voiceless bilabial plosives, a further distinction is made between aspirated and unaspirated [3]ones whereas for voiced bilabial plosives the distinction is between modal-voiced and breathy-voiced ones. The same five-way distinction is also found in dental, alveolo-palatal, retroflex, and velar plosives. In terms of place of articulation, on the other hand, alveolars are the most complex because they involve all manners of articulation except for affricate

Alvelor Plosive

Alveolar plosive is a characteristic of Malayalam and Assamees. Phonetic realization of alvelor plosive of English is different from that of Malayalam. Alevelor plosive have a good amount of friction at the time of release which sets them apart from their retroflex counter parts. It is the sole member of alvelor plosive, it is rather unusual because it is both aspirated and palatalized with no other counterpart in alvelor plosive category[4].

Speech recognition analysis

For conducting speech recognition performance of unique phonemes, we have conducted a special procedure. Initially, the words are recorded with carrier words like " njaan /the word / enn'u paranji' (i spoke the word /word/). This is to nullify the domination of language model in the speech recognition performance. Then the minimal pair counterpart also recorded with the same carrier words. Then we analyse the speech recognition performance . Here we can make a clear decision of whether these two words have been misclassified or not , thereby can make conclusion about the behavior of the phonemes under study.

 For the study of these phoneme, we have selected minimal pairs of words and speech recognition performance of these minimal pairs are also been reported. As the preliminary step towards analysis, special database (called minimal pair) have been designed The second step is to collect speech corpus. The third step is speech Recognition analysis.

Design of database :   We have selected 4 minimal pairs for the study. The following are the words that we have designed.

With retroflex plosive
- പാറ്റ , പാട്ട (/ paat't'a/- cockroach ,   /paatta/ - tin pot)
- പറ്റ് , പട്ട് (pat't'u'/ - buying on credit , /pattu' / - silk)
- കറ്റ , കട്ട   (/kat't'a/ - bundle as of grass , straw etc, /katta/- mass of clay or soil)
- കുറ്റി , കുട്ടി (/kutt't'i / - stump ,   ,/ kutti / - child)

II.   SPEECH RECOGNITION PERFORMANCE

Alveolar Plosive vs Retroflex plosive : The target tokens (patta, katta , paat't'a and kat't'a) were ready by 25 speakers, so that there are a total 100 tokens in the database. For training , we have selected data of 20 speakers . i.e 80 tokens and the remaining 20 tokens were kept for testing. The system has been tested with 20 tokens. Confusion matrix have been drawn for the target phonemes   as in table   minimal pairs evaluated for those tokens and a confusion matrix have been plotted as in table 1.1   It is clear that a alvelor plosive confuses 40% with retroflex plosive in speech recognition.

Table 1.1 confusion matrix of speech recognition performance- tta vs t't'a

	tta	t't'a	total
tta	10	0	10
t't'a	4	6	10
total	14	6	20

III.   CONCLUSION

It is clear that a alvelor plosive confuses 40% with retroflex plosive in speech recognition.From the speech Recognition performance of alveolar plosive and   retroflex plosive, it can be concluded that both has similar auditory stuff. Hence alvelor polosive which is a unique phoneme of the language, has acoustic and auditory properties similar to retroflex plosive which make them confusing in speech recognition performance.

REFERENCES
[1].   Bhaskararao P., "Salient phonetic features of Indian languages", Sadhana, 36(5), pp. 587-599, Oct. 2011.
[2].   Punnoose, R. (2010). *An Auditory and Acoustic Study of Liquids in Malayalam.* Ph.D. Thesis, Newcastle University, Newcastle, UK
[3].   Furui, S., "50 Years of Progress in Speech and Speaker Recognition Research Identification", In ECTI Transformations on Computer and Information Technology, vol. 1, no. 2, 2003
[4].   Rabiner, L. Juang, B. H., Yegnanarayana, B., "Fundamentals of Speech Recognition", Pearson Publishers, 2010.