



Automatic Speech Recognition of Plosive Phonetic Class Words in Malayalam Language

Cini Kurian

Associate Professor, Al-Ameen College, Edathala, Aluva, cinikurian@gmail.com

Abstract: Automatic speech recognition has tremendous potential in Indian scenario. Although literacy rate of India is above 65%, less than 6% of India's total population uses English for communication. Since the internet has become universal, common man now mainly depend the same for any sort of information and communication. Therefore it is imperative that the about 95% of our population cannot enjoy the benefits of this internet revolution. If these information is available in local languages, India could also be benefited by this technology revolution and could stand along with developed countries.

Keywords: Automatic speech recognition , Malayalam, Plosive class words

1. Introduction

In this paper, we discuss the results of the phonetic class wise speech recognition performance of Malayalam language. The words have been collected in such a way that, each word contains at least any of the phonemes of each class. Maximum care has been taken in the selection of words, so that all phonemes of the particular class occurred in all the word positions.

Commonly, in the selection of text corpus for speech recognition, some phonemes never likely to occur and the words which includes such phonemes will never be recognized properly[1]. Therefore special care to be taken to include all phonemes of a language in maximum word positions(Start, end , middle). Our ultimate aim in this work is to develop a absolute speech recognizer for Malayalam Language[2]. Towards this objective we have chosen words in the above manner to include all phonemes of the language in all co-articulatory positions[3]. Thus this work will resolve co-articulation problem to some extent which is a superior challenge for a speech recognizer.

2. Phonetic classes and phones in Malayalam language

Malayalam language has 5 category of phonetic classes, they include stops (aspirated and un-aspirated), Nasals, fricatives, lateral and rhotic[4]. Table 1 list the phones in each category. Each phoneme classes are explained in detailed in the following section.

Table 1: List of Phonetic classes and phonemes in Malayalam

Phoneme classes	Malayalam Phonemes
Unvoiced unaspirated stops	പ p , ബ b, ത t , ദ dh , ട t' , ട d , ഡ ട' , ച ch , ജ j , ക k , ഗ g
voiced aspirated stop	ഫ ph ,ഭ bh, ള th, ഡ dh , റ d'h , റ ട' , റ ch' , റ jh, റ kh, റ gh
Nasals	മ m, ന n, ന്ന n1 , ന്ന' n' , ന്ന ñ , ണ ñ



Frivatives	ഫ f , സ s , ഷ s'h , ശ sh , ഹ h
Lateral	ല l , ല്ല l' , ഴ zh
Rhotic	ര r' , റ r
semi vowels	വ v , യ y

3. Stop class

Stop consonants represent one of the broad categories of phones in all major languages. The production of a stop involves a complete closure of the oral cavity followed by release in the form of noise burst. The stop consonants are differentiated from each other in terms of the manner of articulation (whether voiced and/or aspirated) and the place of articulation. Stops are the categories having maximum number of phones in Malayalam language[5]. They occur in aspirated, un aspirated, voiced and unvoiced forms. In Malayalam there are sixteen stops. Voiced aspirated stops are absent in Malayalam. All voiceless stops have corresponding voiceless aspirated varieties also.

4. Database Design

We have collected words in such a way that each phoneme should occur in initial or, medial or final positions in the word. In all the positions the phonemes are succeeded by the maximum possible vowels. Position wise listing of number of words of each phoneme is detailed in table 1. As an example for the phoneme /p/, in the initial position, we have the words as shown below.

- /paka/ - succeeding vowel 'a'
- /paata/ - succeeding vowel 'aa'
- /piravi/ - succeeding vowel 'i'
- /piili/ - succeeding vowel 'i'
- /puli/ - succeeding vowel 'u'
- /puur'am' / - succeeding vowel 'uu'
- /pettennu'/ - succeeding vowel 'e'
- /pe'tta/ - succeeding vowel 'e''
- /pain laappiL/ - succeeding vowel 'ai'
- /pot'i / - succeeding vowel 'o'
- /po'kan lam'/ - succeeding vowel 'oo'
- /paur'usham'/ - succeeding vowel 'au'

Table 2: Number of words of each phoneme (position wise listing)

	Number of words with phonemes in			total
	Initial position	Medial position	Final position	
Labial voiced unaspirated stop പ	14	12		26
Labial unvoiced unaspirated stop പ്ല	10	10		20
dental voiced unaspirated stop റ	11	9		20
dental unvoiced unaspirated stop റ്ല	10	10		20
alveolar voiced unaspirated stop റ്റ				0



retroflex voiced unaspirated stop ᱥ	1	16		17
retroflex unvoiced unaspirated stop ᱦ	5	10		15
palatal voiced unaspirated stop ᱠ	12	8		20
palatal unvoiced unaspirated stop ᱡ	11	10		21
velar voiced unaspirated stop ᱢ	13	10		23
velar unvoiced unaspirated stop ᱣ	10	9		19
Total				201

5. Design and Development

Speech recognizer for all the above classes of words separately. Semi continuous, Context dependent tied state HMM's with 3 state per HMM[6] and 8 Gaussian[7] per state were used for modeling. MFCC[8] were used for feature extraction and trigram models used for language modeling. In each phonetic class wise recognizer, speech corpus contain 25 speakers' data, out of which training is performed by 20 speakers' data and testing by 5 speakers' data. The experiment was conducted using 5 fold validation test as explained in section and the results were analyzed using the performance metric WER[9] using sclite from NIST[10].

6. Results and discussions of Stop class words (Words having stop class phonemes)

The words in this category consists of 205 words which include 26 labial voiced stop words 20 labial unvoiced stop words , 20 dental voiced stop words, 20 dental unvoiced stop words, 23 velar voiced stop words, 19 velar unvoiced stop words, 21 retroflex unvoiced stop words, 20 palatal voice stop words, and 22 palatal unvoiced stop words.

The speech performance results of these words in training and testing modules are as shown in table 3. The average accuracy obtained for test data is 82%.

Table 3: Speech Recognition results words having stop class phonemes

Sl.No	Training %	Testing %
1	93	80.15
2	94.28	83.78
3	94.12	83.5
4	93.14	80.79
Average	93.635	82.055

7. Summary

In this paper, we have presented our effort of developing recognizers for different categories of words. Here we have classified the words by phonetic class. This work can be treated as a contribution to the domain of Isolated word recognizer for Malayalam language, since a total of 201 words have been identified and recognized. This corpus can be claimed to be a phonetically balanced corpus and hence the recognizer would be appropriate for any type of applications as these words include all phonemes of the language in all word positions.



References

- [1] V. Vapnik. Statistical Learning Theory. Wiley, NY, 1998.
- [2] C. J.C. Burges, "A tutorial on support vector machines for pattern recognition," knowledge Discovery Data Mining, vol. 2, no. 2, pp.121- 167, 1998.
- [3] A. Ganapathiraju, . Hamaker, and. Picone, "Hybrid SVMHMM architectures for speech recognition," in Proc. ICSLP, Beijing, 2000
- [4] Alexandros Karatzoglou, David Meyer,Kurt Hornik , Support Vector Machines in R Journal of Statistical software April 2006, Volume 15,Issue 9 (5)
- [5] L.E. Baum and T.Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," Ann.Math.Stat.,vol.37,pp.1554-1563, 1966
- [6] Levinson, S. E., Rabiner, L. R Sondhi, M. M., (1983). Speaker Independent Isolated Digit Recognition Using Hidden Markov Model, In Proceedings of ICASSP, pp.1049-1052.
- [7] L.,R Rabiner, "A tutorial on Hidden Markov model and selected application in speech recognition", Pro.IEEE,7(2):257-286, February 1998
- [8] Hwang, M. Y. (1993). Sub-phonetic acoustic modeling for speaker-independent continuous speech recognition. Ph.D. Thesis. Carnegie Mellon University.
- [9] C.H Lee, L.R Rabinar , R. Pieraccini, and J.G Wilpon , " Acoustic Modelling for Large Vocabulary speech Recognition ", Computer speech and Language , 4:127-165,1990 Vocabulary speech Recognition", Computer speech and Language , 4:127-165,1990
- [10] Young, S., Evermann, G., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V. and Woodland, P. (2002). The HTK Book (for HTK Version 3.2).Microsoft Corporation and Cambridge University Engineering Department, England.