International Journal of Computer Applications Technology and Research
Volume 7–Issue 05, 200-202, 2018, ISSN:-2319–8656

Automatic Speech Recognition of Malayalam Language Nasal Class Phonemes

Cini Kurian
Associate Professor, Al-Ameen College
Edathala, Aluva

Abstract: Speech recognition applications are becoming common and useful in this day and age as many of the modern devices are designed and produced user-friendly for the convenience of general public. Speaking/communicating directly with the machine to achieve desired objectives make usage of modern devices easier and convenient. Although may interactive software applications are available, the use of these applications is limited due to language barriers. Hence development of speech recognition systems in local languages will help anyone to make use of this technological advancement. In this paper we discuss the results of the Nasal phonetic class wise speech recognition performance of Malayalam language

Keywords: Automatic speech recognition, Malayalam, Nasal class words

1. INTRODUCTION

Speech recognition technology is highly dependent on the spoken language. Phonetic and acoustic features of each phoneme in a language is the prime factor for speech recognition technology. In speech technology and phonetics , speech signal is the basic material for study and analysis. 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. But since in most of the languages, phonemes and letters have a varying degree of correspondence, the allophones are mapped into its higher classes called phonemes. Indian languages have more direct relation between sounds and letters[1]. 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[2]. Phonetic realization of all the phonemes will not be the same across languages. Some language has some peculiar phonemes and some phonemes has different realization. Malayalam, a language that has not been the subject of much research in the speech recognition as well as phonetic and phonological literature[3]. So in this work , we have tried to explore the speech recognition of Nasal class phoneme of Malayalam language.

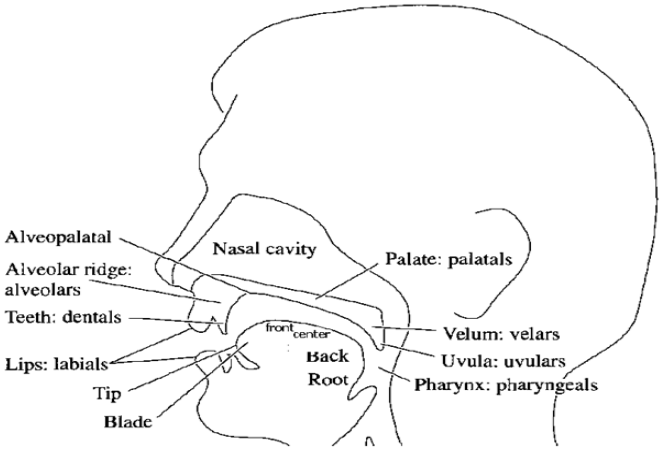
Commonly, in the selection of text corpus for speech recognition, some phonemes never likely to occur and the words which includes such phonemes will never be recognized properly[4]. Therefore special care to be taken to include all phonemes of a language in maximum word positions(Start, end , middle). Our ultimate aim in this work is to develop a absolute speech recognizer for Malayalam Language[5]. Towards this objective we have chosen words in the above manner to include all phonemes of the language in all co-articulatory positions[6]. Thus this work will resolve co-articulation problem to some extent which is a superior challenge for a speech recognizer

2. PLACE OF ARTICULATION AND PHONETIC CHART

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 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[7]. Phonetic chart as presented by Kumari, 1972 [8] for Malayalam language is given in table 1.

Table 1. List of Phonetic classes and phonemes in Malayalam

Phonome classess	Malayalam Phonemes
Unvoiced unaspirated stops	പ p , ബ b, ത t , ദ dh , ഠ ṭ , ഡ ḍ , ച ch , ജ j , ക k , ഗ g
voiced aspirated stop	ഫ ph ,ഭ bh, ഥ th,ധ dh , ഢ ḍh , ഝ jh, ഖ kh, ഘ gh
Nasals	മ m, ന n, ണ n1 , ഞ ñ , ങ ṅ
Frivatives	ഫ f , സ s , ഷ ṣ h , ശ sh , ഹ h
Lateral	ല l , ള ḷ , ഴ zh
Rhotic	ര r' , റ r
semi vowels	വ v , യ y



For all speech sounds, the basic source of power is the respiratory system pushing air out of the lungs. Sounds produced when the vocal cords are vibrating are said to be voiced, where the sound produced when the vocal cords are apart are said to be voiceless[8]. The shape and size of the vocal tract is a very important factor in the production of speech. The parts of the vocal tract such as the tongue and the lips that can be used to form sounds are called articulators. The movements of the tongue and lips interacting with the roof of the mouth(palate) and the pharynx are part of the articulatory process[9].

3. NASAL CLASS

In Malayalam there are six nasals. The allophonic distribution of these six nasals are given below. We produce these nasal phonemes by lowering the velum to allow air to flow through the nasal cavity. The three nasals listed above are consonants which involve complete obstruction of the oral cavity. The location of the obstruction determines the place of articulation of the nasal: the lips for /m/, the tip or blade of the tongue at the alveolar ridge for /n/, and the back of the tongue at the velum for /n1/[10]
The six nasal class of phonemes and their allophonic distribution are shown below:

[m] Bilabial nasal occurs initially, medially and finally
e.g. /mashi/ ‘ink’ , /umi/ ‘paddy husk’ , /maram/ ‘tree’

[n1] Dental nasal occurs initially independently and medially in clusters
E.q. /innu' / ‘today’ , /nayam'/ ‘policy’ ,/ nruttam' / ‘dance , /cantam'/ ‘beauty’

[n] Alveolar nasal occurs initially before [y], medially and finally.
E.g. /n1ayam' / ‘justice’ , /vin1ayam/ ‘obedience’ , /miin/ ‘fish’

[N] Retroflex nasal. occurs medially and finally.
e.g. /pan'am'/ ‘money’ , /juu:N1/ ‘June’

[ň] palatal nasal occurs initially and medially.
E.g. /ñaaN/ ‘I’ , / maňňu'>/ ‘dew’

[ŋ] velar nasal.occurs medially with length or in clusters.
E.g. /maaṅṅa/ ‘mango’ , /ta ṅ ṅi/ ‘stayed’

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

Table 2 Number of words of each phoneme

	Number of words with phonomes in			total
	Initial position	Medial position	Final position	
Labial nasal ൺ	12	8		20
Dental nasal ൻ1		14	4	18
alveloar nasal ൻ		16		16
retroflex nasal ൺ		18	3	21
Total				75

5. DESIGN AND DEVELOPMENT

We have developed speech recognizer for all the above classes of words separately. Semi continuous, Context dependent tied state HMM's with 3 state per HMM and 8 Gaussian per state were used for modeling. MFCC 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 and the results were analyzed using the performance metric WER[11] using sclite from NIST.

6. RESULTS AND DISCUSSIONS OF NASAL CLASS WORDS

Nasal class words are designed to include total of 74 words which include 20 labial nasal words, 20 dental nasal words, 14 alveolar nasals and 20 retroflex nasals .The classification performance of testing and training modules in each of the

four experiments (fivefold validation) is as shown in table 3. The average accuracy obtained from the training data is 97 and that of test data is 72 percentages respectively.

Table 3. Speech recognition of words having nasal class phonemes

Sl.No	Training %	Testing %
1	98.65	70.61
2	97	75.97
3	96	68.75
4	96.65	69.25
Average	97.15	71.96

7. SUMMARY

In this chapter we present and discuss the results of the experiments we have conducted for Malayalam language speech recognizer for Nasal class words. Data base for the phoneme class words have been designed and its speech recognition performance is calculated. The average accuracy obtained from the test data is 72 percentage.

8. REFERENCES

[1] Furui, S., “50 Years of Progress in Speech and Speaker Recognition Research Identification”, In ECTI Transformations n Computer and Information Technology, vol. 1, no. 2, 2003

[2] Sorin Dusan and Larry R. Rabiner, “On integrating insights from human speech perception into automatic speech recognition,” in Proceedings of INTERSPEECH 2005, Lisbon, 2005.

[3] HILL, D. R. (1971). Man-machine interaction using speech. In Advances in Computers, 11. Eds F. L. Alt, M. Rubinoff & M. C. Yovitts, pp. 165-230. New York: Academic Press.

[4]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

[5] Balaji. V., K. Rajamohan, R. Rajasekarapandy, S. Senthilkumaran,"Towards a knowledge system for sustainable food ecurity: The information village experiment in Pondicherry," in IT Experience in India : Bridging the Digital Divide, Kenneth Keniston and Deepak Kumar, eds., New Delhi, Sage,2004.

[6] Madhuresh Singhal et al. ‘Developing Information Technology Solutions in Indian Languages: Pros and Cons’. At 1st International CALIBER: Mapping Technology on Libraries and People, 13-15 Feb. Ahmadabad, India, pages 655-666, 2003.

[7] Namboothiri, E.V.N. 2002. Bhashavinjaneeyam. Calicut: Poorna Publications

[8] Punnoose, R. (2010). An Auditory and Acoustic Study of Liquids in Malayalam. Ph.D. Thesis, Newcastle University, Newcastle, UK

[9] G. Doddington, (1989), "Phonetically Sensitive Discriminants for Improved Speech Rec.", Proc. IEEE Int Conf. Acoustics. Speech and Sig. Proc., ICASSP-89, pp. 556-559, Glasgow, Scot- land.

[10] L.,R Rabiner, "A tutorial on Hidden Markov model and selected application in speech recognition" , Pro.IEEE,7(2):257-286, February 1998