Durational Characteristics of Indian Phonemes for Language Discrimination

B. Lakshmi Kanth, Venkatesh Keri, and Kishore S. Prahallad

International Institute of Information Technology, Hyderabad, India – 500032
{lakshmikanth.b,venkateshk}@research.iiit.ac.in,
kishore@iiit.ac.in

Abstract. Speech is the most important and common means of communication. Human beings identify a language by looking at the acoustics and the letter to sound rules (LTS) that govern the language. But pronunciation is governed by the person's exposure to his/her native language. This is a major issue while considering words, especially nouns in Indian languages. In this paper, a new methodology of analyzing phoneme durations for language discrimination is presented. The work has been carried out on a database built with words, mostly nouns, common to Hindi, Tamil and Telugu languages. Durational analysis of phonemes has been carried out on the collected database. Our results show that phoneme durations play a significant role in differentiating Hindi, Telugu and Tamil languages with regard to stop sounds, vowels and nasals.

Keywords: Phoneme, Duration and Statistical Significance.

1 Introduction

Analysis of language discrimination is important in a country like India with many languages. The differences between languages can be used in language identification, speech recognition, speaker recognition and also for building text-to-speech systems. In the past, the analysis on languages has been carried mostly using prosodic and spectral information. Much of the research on languages so far has been on spectral information, mainly using the phonemic features and their alignment. Such systems may perform well in similar acoustic conditions [4]. Indian languages such as Hindi, Tamil and Telugu share a similar phoneme set in the production of speech sounds. In the case of words that are common to Indian languages, the discrimination rate decreases as the phonemic alignment and their sound almost remains constant for a particular phoneme. Therefore, looking at prosodic features like duration, pitch, intensity is a significant key for language discrimination.

In this paper we are looking into durational characteristics of phonemes for language discrimination. Klatt [1] studied the segmental duration of English language and showed that phoneme durations have potential cue in carrying language information. Samudravijaya [2] stated that durational cues are useful in discriminating Hindi language phonemes. Y.K.Muthusamy [3] stated that phonemic transcriptions

are advantageous than arriving at accurate phonetic transcriptions. The study of durational characteristics of languages with similar phoneme set is therefore an interesting task.

In section 2, building the phonemically balanced database is discussed. It includes collection of words, speech recording and manual labeling of speech. In section 3, analysis of language discrimination is discussed with the final conclusion in section 4.

2   Database and Labeling

In performing the phonemic durational analysis, we have jointly prepared a phonemically balanced word database. The words are nouns e.g., person names, place name, etc. of Indian origin and are common to the three Indian languages Hindi, Tamil and Telugu. A total of 440 words have been collected and ensured that each phoneme occurs at least 5 times in the beginning, middle and end of a word in the database.

The speech database was collected with 25 speakers with an average age group of 18 to 25 years having 13 males and 12 females in each language. All the speakers are bilingual in the sense that they have educational background in native language at least upto senior schooling. Speakers were provided with words in their native script and were instructed to speak each word in their respective language with normal rate and intonation into the microphone. The problem with the recording phase is that each speaker of a particular nativity has his own way of pronouncing the words. This will effect the phonemic coverage in each language. In order to avoid the above problem, we made sure that each word has been recorded with same pronunciation in three languages. All these utterances were recorded using Edirol R09 speech recorder at 48 kHz frequency. Also, the high quality speech is digitized at 16 kHz with 16bits per sample in a quite environment to avoid echo effect.

Table 1. The Indian Language Transliteration (ITRANS) symbols of phonemes common to three languages Hindi, Tamil and Telugu

Semi Vowels					Long Vowels					Dipthongs	
a	i	u	e	o	aa	ii	uu	ei	oo	ai	au

Unvoiced Unaspirated	Unvoiced Aspirated	Voiced Unaspirated	Voiced Aspirated	Nasals
k	kh	g	gh	-
ch	chh	j	jh	nj~
t:	t:h	d:	d:h	nd~
t	th	d	dh	n
p	-	b	bh	m

Semi Vowels				Fricatives			
y	r	l	v	sh	shh	s	h

Manual labeling of the recorded data has been carried out for phoneme durational analysis. Figure 1 illustrates the waveform and spectrogram with phoneme labels of word "auku". A total of 43 phonemes were considered for the study as tabulated in Table 1. In the manually labeled speech database, special symbols were used to signify the geminate consonants, and closure and release durations of stop consonants separately which were tabulated in Table 2. For e.g., the word "auku" in Fig 1 is labeled into phonemes au, kcl, k, and u. The singleton stops have same closure and release stop phonemes, the cluster stops have dissimilar closure and release stop phonemes and gemination refers to repetition of same phoneme.

Table 2. Rules followed in manual labeling

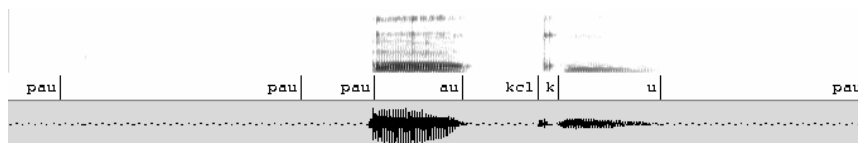|	Consonant	Rule
Singleton stop	k	kcl k
Cluster stop	tk	tcl k
Gemination Stop	kk	k1cl k1
Gemination	nn	n1

Fig. 1. Waveform and spectrogram of word "auku" spoken by a Hindi speaker

3   Analysis

The speech database has words that were intentionally pronounced to have similar sound words. The mean values and standard deviation values of phoneme durations were computed and used to study the durational characteristics in differentiating languages. As shown in Table 3, the broad-phonetic classifications are categorized into short vowels, long vowels, semivowels, diphthongs, nasals and fricatives, and stop sounds. The stop sounds are separately labeled into closure and release phonemes. The following section provides an analysis of the duration of phonemes in the collected word database.

Student's t-test is used in analyzing the significance of durational differences between phonemes. This test is used to calculate the statistical difference as a function of the difference between means relative to the variability [5]. Each of the other columns in Table 3 contains the results of comparison of phoneme durations of any two languages. This was conducted at 99.95% confidence level.

In stop sounds, release durations have more practical significance than closure durations. Stop sounds shows that release durations have statistical significance in the

case of unvoiced unaspirated, unvoiced aspirated and voiced aspirated phonemes. Their durational statistics are tabulated in Table 4. In Hindi, the durations of unvoiced aspirated are twice the durations of unvoiced unaspirated. In Tamil language, durations of unvoiced aspirated are equal to durations of unvoiced unaspirated, this may be because of having single alphabet for the stop sound in the place of articulation. In Telugu, durations of voiced unaspiration are equal to that of voiced aspiration and durations of unvoiced aspiration are twice the durations of unvoiced unaspiration.

The vowels show significant difference between Telugu and Hindi in short vowels and between Tamil and Telugu in long vowels. The vowels helps in differentiating Telugu language, because Hindi and Tamil speakers pronounce both short and long vowels almost at equal duration rates. We found that Hindi has more duration and Tamil has less duration than Telugu in short and long vowels respectively and are tabulated in Table 5.

In the case of nasals, Hindi language has a significant durational characteristic. Hindi phonemes m, n, and n1, have more duration and nd~, nj~ have lesser duration compared to other two languages. This is more of a practical significance in differentiating Hindi language and are tabulated in Table 6.

From the above analysis, Telugu can be discriminated using vowels and Hindi can be discriminated using nasals. Along with the help of singleton stop phoneme durations, the three languages can be discriminated.

Table 3. Table shows the phoneme durations that are statistically significant at 99.95% confidence level between languages Hindi, Tamil and Telugu. "S" represents the significance between two languages.

Broad Phonetic Classification		Hindi-Tamil	Tamil-Telugu	Telugu-Hindi
Short Vowels				S
Long Vowels			S	
Semi Vowels				
Unvoiced Unaspirated	Closure			
	Release	S	S	S
Unvoiced Aspirated	Closure			
	Release	S	S	S
Voiced Unaspirated	Closure			
	Release			
Voiced Aspirated	Closure			
	Release	S	S	S
Diphthongs				S
Nasals		S		S
Fricatives			S	

Table 4. Mean and standard deviations, std (in msec) of release durations of singleton stop sounds in three languages Hindi, Tamil and Telugu languages

		Hindi		Tamil		Telugu	
	Singleton Stops	Mean	Std	Mean	Std	Mean	Std
Unvoiced Unaspirated	k	34.60	11.25	37.72	11.99	39.36	11.91
	t:	20.27	6.02	21.58	5.73	20.52	5.96
	t	25.89	7.90	28.27	8.67	27.98	8.62
	p	23.00	8.32	26.60	9.74	26.69	10.46
Unvoiced Aspirated	kh	76.67	22.22	37.80	10.28	54.94	22.70
	t:h	69.40	24.73	25.26	8.04	43.57	29.83
	th	60.77	20.78	30.36	7.72	40.89	19.45
Voiced Unaspirated	g	60.73	26.21	60.84	26.60	54.93	20.88
	d:	46.46	31.80	43.27	28.88	37.60	25.24
	d	63.73	32.60	58.87	32.70	55.23	23.77
	b	63.18	27.59	65.44	26.49	59.75	21.25
Voiced Aspirated	gh	62.97	30.92	63.94	35.00	53.97	20.01
	d:h	59.35	33.59	64.16	34.53	56.23	28.18
	dh	60.89	28.60	57.05	27.25	52.11	21.96
	bh	60.69	22.70	69.77	21.95	61.85	37.35

Table 5. Mean and standard deviations,std (in msec) of short and long vowels between Hindi and Telugu and between Telugu and Tamil languages respectively

Phone	Hindi		Telugu	
	Mean	Std	Mean	Std
a	79.080	31.137	76.776	23.936
e	109.439	43.779	92.286	38.739
i	106.740	59.646	99.948	53.136
o	103.772	29.205	75.068	22.171
u	85.850	41.814	78.676	39.406

Phone	Telugu		Tamil	
	Mean	Std	Mean	Std
aa	156.671	40.531	164.670	42.208
ei	148.157	37.818	156.548	37.861
ii	122.819	37.657	131.766	35.809
oo	135.856	33.975	145.210	33.652
uu	125.987	42.452	137.889	41.624

Table 6. Mean and standard deviations,std (in msec) of nasal sounds

Phone	Hindi		Telugu		Tamil	
	Mean	Std	Mean	Std	Mean	Std
m	82.59	27.94	77.70	24.10	78.69	27.24
n	82.44	31.82	76.10	31.03	76.59	31.82
n1	130.89	27.37	124.98	23.87	117.37	18.78
nd~	67.62	38.06	81.32	34.34	80.70	35.20
nj~	74.11	29.35	84.34	35.56	79.90	31.23

4 Conclusion

A study on the analysis of language discrimination on phonemically balanced speech database is presented. With intentionally spelled words having same pronunciation, we find major differences in the prosodic level. From the analysis, Telugu can be discriminated using vowels and Hindi can be discriminated using nasals. Along with the help of singleton stop phoneme durations, the three languages can be discriminated. Hence, phoneme duration holds key information for language discrimination.

References

1. Klatt, D.H.: Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. J. Acoust. Soc. Am. 59(5), 1208–1221 (1976)
2. Samudravijaya, K.: Durational characteristics of Hindi phonemes in Continuous Speech. Technical Report, TIFR (April 2003)
3. Muthuswamy, Y.K.: A segmental approach to Automatic Language Identification, Ph.D. Thesis, Oregon Graduate Institute (1993)
4. Zissman, M.A.: Comparison of four approaches to automatic language identification of telephone speech. IEEE Transactions on Speech and Audio Processing 4(1), 31–44 (1996)
5. Douby, S., Weardon, S., Chilko, D.: Statistics for research. John Wiley and Sons, Inc., Hoboken (2004)