

Durational Characteristics of Allophonic Variations in Malayalam Consonant Phonemes

Vivek P¹, Sandesh E PA², Lajish V L³

Department of computer science

University of Calicut

Kerala, India

(¹vivek_dcs@uoc.ac.in, ²sandeshepa@gmail.com, ³lajish@uoc.ac.in)

Abstract—Phonemes, defined as the smallest distinctive sound units in a language, are considered to be the basic unit for speech. But the properties of phonemes exhibit wide variations based on its position in the word and context. In Malayalam, phonemes are further categorised in to allophones based on the positional and contextual variability, i.e. the contextual and positional variability is encoded in the allophone characterisation of Malayalam, a major Dravidian language spoken by around 38 million people. A well-defined allophone formation rule set exists for Malayalam. In this work consonant allophones in Malayalam are identified and listed. An allophone centric durational model is developed as part of this work. A detailed analysis is performed which reveals the basic durational characteristics of Malayalam consonant phonemes.

Keywords—Phoneme; Allophone; Consonant; Duration; Speech Recognition; Speech Synthesis.

I. INTRODUCTION

In the widely used 35 languages of the world one third is from India. Malayalam, a south Indian language spoken by around 35 million people, is a classical Indian language and it is the official language of Kerala [1]. Malayalam, a low resource language, needs extensive studies to develop automatic language processing tools addressing its inherent peculiarities.

Speech being the most natural method of communication, interacting with a machine through speech is the most explored area in developing man – machine interfaces. Automatic speech recognition and synthesis are the main speech processing tasks. Effectiveness of speech processing systems greatly depends on results from language specific explorations on many features. Duration and its modelling is an important cue affecting the intelligibility and naturalness of synthesised speech.

Phonetic identity of current segment, surrounding segments, positioning within the word and within the sentence, emotional state of the speaker *etc.* are factors affecting the duration of a segment [2]. It is possible to capture most of positional and contextual variability from the textual representation of the sentence. A phoneme can appear in the start, middle and end of a word creating positional variation. The change in duration due to the effect of surrounding speech

units is called contextual variability. The effect due to neighboring phonemes and position are generally known as co-articulation effects. Co-articulation effects in Malayalam are modelled by Malayalam linguists as allophonic characterisations for each phoneme. A well-defined allophone formation rule set exists for Malayalam [3]. So a durational model accommodating co-articulation effects due to positional and contextual variability of phonemes can be developed by understanding the durational pattern of allophones. Such rule based approach for phoneme duration modelling has been reported for many languages. Rule based segmental duration approaches have started with the work of Klatt for English [4] which is the basis for many Text to Speech systems. Rule based duration models has been established for many languages including French, Brazilian and Portuguese [5-6]. A combined rule based and statistics based prosody modelling is also applied for concatenative speech synthesis in Tamil and Hindi [7].

Vivek P *et. al.* studied the durational characteristics of allophonic Variations in Malayalam vowels in detail [8]. In this work, consonant allophones in Malayalam are identified and listed with its allophonic variations. Section II describes the properties of Malayalam consonant phones and allophones. Section III discuss the process of finding the extensive rule set for the formation of Malayalam consonant allophones. Section IV introduces the TEMU Malayalam phonetic archive. Section V discusses durational properties of Malayalam vowel allophones and section VI concludes the work.

II. MALAYALAM PHONEME SET AND ALLOPHONES

A phone refer to the instances of phonemes in the actual utterances. The smallest meaningful distinctive sound unit in a language is called a phoneme. The phoneme set varies considerably from one language to another. International Phonetic Alphabet (IPA) defines around 150 phones among all languages. American English has around 40 phonemes. Malayalam has 11 vowel phonemes, 2 diphthongs, and 37 consonant phonemes together constitute a 50 member phonemeset [9]. The following section describes the characteristics of phonetic inventory of Malayalam language and its allophonic variations in detail followed by the structure of the Malayalam phonetic archive used in this study.

A. Malayalam consonant phones

A consonant is a speech sound that is articulated with the complete or partial closure of the vocal tract [10]. The classification problem of consonants is more challenging than that of vowels. There are more parameters of contrast with the neighbors on consonant classification. Consonants can be broadly categorized into plosive (stop consonant), nasal, fricative, flapped, lateral, approximant and glide. Malayalam has 38 consonants in which 21 of them are plosives. Plosives are again classified based on voice and aspiration. The classification of Malayalam plosives are given in table I and the consonants other than plosives are listed in table II.

TABLE I. CLASSIFICATION OF MALAYALAM PLOSIVES

	Voice	Bilabial	Dental	Alveolar	Retroflex	Palatal	Velar
Plosive	Voiceless unspirated	P പ	t ത	ɾ റ	ʈ റ	c ച	k ക
	Aspirated	p ^h പ്	t ^h ത്		ʈ ^h റ്	c ^h ച്	k ^h ക്
	Voiced Unspirated	b ബ	d ഡ		ɖ ഡ	ɟ ജ	g ഴ
	Aspirated	b ^h ബ്	d ^h ഡ്		ɖ ^h ഡ്	ɟ ^h ജ്	g ^h ഴ്

TABLE II. CLASSIFICATION OF MALAYALAM CONSONANTS OTHER THAN PLOSIVES

	Bilabial	Labiodentals	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Nasal	m മ		ɳ ണ	n ന	ɳ ണ	ɲ ണ	ŋ ണ	
Fricative				s ശ	ʂ ഷ	ʃ ഴ		h ഴ
Trill/flapped				r റ				
Lateral				l ല	ɭ ല			
Approximant					ɻ ഴ			
Glide		v വ				y യ		

III. AN EXTENSIVE RULE SET FOR THE MALAYALAM CONSONANT ALLOPHONE FORMATION

The variation in the duration of a phoneme can be attributed to many factors. Some factors such as contextual variability can be detected from the text while some others such as dialect cannot be detected from the text. This work addresses the durational variability of vowel allophones in Malayalam. In their studies, Asher and V.R. Prabodhachandran Nair described the rules of Malayalam allophone formation [11,3]. They have proposed linguistic descriptions for defining the allophones of each Malayalam phones. It reveals that most such findings can be converted to position and neighborhood-based rule set. Thus certain rule set for Malayalam consonant allophones based on position and neighboring information's are created. The following section describes the rule set for the formation of allophones in detail.

A. Rule set for Malayalam consonant allophones based on the position and neighbourhood information

A rule set for the formation of consonant allophones in Malayalam are also constructed based on the position and neighbouring information. The list of Malayalam consonants with its allophones is shown in table III. It can be seen that there are nineteen consonants with only one allophone and few allophones have more than one rule for its formation. For example, the allophone [v] of the vowel റ് /v/ has three rules for the formation based on its position and neighbourhood. In this case the allophones may occur in (a) word initial position (b) medial position (in clusters where [w] does not appear) (c) mostly short and rarely long intervocalic positions.

In another case, the consonant phone മ [m] has 4 allophones in Malayalam. The allophone [m^h] occurs before velar fricative. The next allophone [M] occurs in consonant clusters when preceded by alveolar flap. The third allophone [m] is characterized by the presence of labio-dental continuant before it and the final allophone [m] (bilabial nasal) occurs elsewhere other than the above three allophones.

TABLE III. POSITION AND NEIGHBOURHOOD BASED RULE SET FOR MALAYALAM CONSONANT ALLOPHONES

Sl. No.	Phoneme	Allophone	Rule
1	പ [P]	[p]	Metadata: Voiceless tense bilabial stop contoid. Initially.
		[β]	Metadata :Voiced bilabial approximant . Voiced bilabial stop contoid with more lax quality . Intervocalically
		[b]	Metadata :Slightly voiced bilabial stop . Slightly voiced bilabial stop contoid In medial nasal – plosive cluster]
		[P]	Metadata: Voiceless most tense bilabial stop contoid. Medially singly or in a cluster except when preceded by nasal.

2	फ [ph]	[ph]	Metadata: -Voiceless aspirated bilabial stop . Occurs initially and medially in Sanskrit loans	8	थ [th]	[th]	Metadata: Voiceless aspirated dental stop. Voiceless aspirated lamino-dental stop. . Medially in Sanskrit loans.
3	ब [b]	[B]	Metadata: Voiced tense bilabial stop contoid. In consonant clusters.	9	ड [d]	[d]	Metadata: voiced tense dental stop contoid . medially in consonant clusters
		[b]	Metadata: Voiced lax bilabial stop contoid. . occurs initially and intervocally.			[d]	Metadata: voiced dental stop contoid with lax quality. . Initially, intervocalic, and in clusters.
4	भ [bh]	[bh]	Metadata: Voiced aspirated labio-labial stop. Occurs initially and medially in Sanskrit loans.	10	ध [dh]	[dh]	Metadata: Voiced aspirated lamio dental stop . Occurs initially and medially in Sanskrit loans.
5	म [m]	[ṃh]	Metadata: Voiceless bilabial nasal contoid . 1. Before velar fricative	11	न [n]	[ṃ]	Metadata: More tense dental nasal contoid. 1. Preceded by alveolar flap.
		[M]	Metadata : More tense bilabial nasal contoid . 1 In consonant clusters when preceded by alveolar flap			[n]	Metadata: Less tense dental nasal contoid. 1 Elsewhere, i.e. short initially and before the other dental consonants 2. Long in intervocalic position
		[m]	Metadata : Labio – dental nasal . 1 Before labio – dental continuant				
		[m]	Metadata : labial nasal .. Elsewhere				
6	व [v]	[w]	Metadata: Voiced bilabial continuant. Preceded by consonants except flapped in consonant clusters.	12	र [r]	[d]	Metadata: Voiced alveolar stop. Voiced lax alveolar stop contoid. 1 In a medial homorganic nasal stop [In such sequence the -n is a stem-final consonant and [r] is the first consonant of the genitive case suffix] 2. After a nasal
		[v]	Metadata: Voiced labiodental continuant. 1 Initially 2 In the clusters in medial position, where [w] does not appear. 3. Mostly short and rarely long in intervocalic position.			[t]	Metadata: Voiceless tense alveolar stop contoid. Voiceless apico-alveolar stop with, for some speakers, a slight palatal quality and/ or a hint of affrication. 1 As identical consonant cluster in intervocalic position .. 2 Medially when it is long ..
7	त [t]	[t]	Metadata: Voiceless lamino-dental stop., Voiceless tense dental stop contoid , Occurs initially.	13	न [n]	[nh]	Metadata :Voiceless alveolar nasal contoid. When preceded by velar fricative /h/
		[t[i]	Metadata: Voiceless most tense dental stop contoid .In clusters except when not preceded by nasals and [j], medially when geminated.			[n]	Metadata: Voiced alveolar nasal contoid, elsewhere. . 1. Word final position 2. Short before other alveolar consonant 3 Short or long in intervocalic position
		[ð]	Metadata: Voiced lamino dental approximant possibly with slight friction. More voiced dental stop contoid with more lax quality. Intervocally or preceded by [j]				
		[d]	Metadata: voiced lamino-dental stop. Slightly voiced dental stop contoid with lax quality. . Medially preceded by nasal.				
14	स [s]	[s]	Metadata: voiceless apico alveolar fricative. Occurs initial medial and final position. In the case of final /s/ there is an alternative pronunciation with an added enunciative vowel'., Voiceless denti – alveolar sibilant slit fricative .. Singly in initial position and in clusters, long intervocally and medially				

15. a	ര [r]	[r]	Metadata: voiced apico denti alveolar tap . 1. Word initially 2. The second consonant in some initial consonant sequence. 3. Intervocally. 4. In a number of medial clusters. Voiced palatalized denti alveolar flap contoid . 1. Word initially 2. Intervocally 3. before [j] 4. after [b] [d] or [g]	21	ണ [ɳ]	[ɳ]	Metadata: Voiced sub lamino – palatal (retroflex) nasal. Retroflex nasal contoid. Occurs medially ie, 1. Intervocally 2. Medially as a geminate 3. In the following medial clusters ɳ, ɳɳ, ɳj, ɳɳ. Short in word final position [tu: .]
15. b	റ [r]	[r]	Metadata: - Voiced apico-alveolar tap or trill . 1. Word initially 2. Second consonant in initial consonant sequences 3. Intervocally 4. In a number of medial sequences Voiced velorized alveolar flap. 1. Rarely in initial position 2. Intervocally 3. Finally 4. Followed by consonants except [j] 5. After consonants except [b] [d] [g]	22	ഷ [ʃ]	[ʃ]	Metadata: Voiceless apico alveolar fricative . Voiceless retroflex more long, tense sibilant groove fricative . In initial, medial and final positions in loan words.
16	ല/ൽ [l]	[l]	Metadata: Voiced apico – alveolar lateral . 1 Word initially 2 The second element in some word initial clusters 3. Intervocally 4 Medially as a geminate consonant. 5. Word-finally. Voiced frictionless alveolar lateral contoid . 1. Rarely in word initial position 2. Short or long intervocally 3. Finally 4. In clusters	23	ള/ൾ [ɭ]	[ɭ]	Metadata - Voiced sublamino palatal (retroflex) lateral . 1 The second element in some word initial clusters in loans 2. Intervocali ally 3. Medially as a geminate consonant. 4. Finally Voiced frictionless retroflex lateral contoid . 1. Finally 2. In clusters 3 Short or long intervocally. 4. Rarely in initial position.
17	ട [t]	[d]	Metadata: Slightly voiced and laxed retroflex stop contoid. . Voiced sub lamino postalveolar (Retroflex) . After homorganic nasal.	24	ഴ [z]	[z]	Metadata : Voiced sublamino palatal approximant. 1 Intervocally 2. First element in medial consonant clusters 3. Finally (in which, there is alternative with vocalic release) Voiced retroflex continuant. 1 Medially, intervocally 2. In consonant clusters. VRP posits that the concerned Malayalam sound is without even a trace of friction and employs a new symbol [y] instead of (z).
		[t]	Metadata: More voiced and lax retroflex plosive contoid intervocally. Voiced sublamino post alveolar flap. .				
		[ʈ]	Metadata: Voiceless tense retroflex plosive contoid. voiceless sublamino post alveolar (Retroflex). In word initial position in loan words.				
		[ʈ]	Metadata: Voiceless retroflex plosive contoid with more tense quality. . In consonant cluster; not preceded by nasal.				
18	ത [th]	[tʰ]	Metadata: Voiceless aspirated tense retroflex. . Intervocally and preceded by nasal	25	ച [c]	[c]	-
		[Tʰ]	Metadata: Voiceless aspirated more tense retroflex. . Elsewhere; not after nasal in clusters.			[c̥]	Metadata : Voiceless aspirated tense velar plosive; initially .
19	ഡ [d]	[d]	Metadata: Voiced sublamino postal velar stop . Voiced retroflex stop. Occurs initially and medially. Both intervocally and in the sequence / d/ . lax when short and tense when long .	26	ഛ [ch]	[j]	Metadata: Voiced lamino – palato- alveolar stop, slightly affricated. Occurring medially after /p /.
		[d]	Metadata: Voiced palatal affricate with maximum tense quality. In clusters; not preceded by nasals. .			[C]	Metadata: Voiced palatal affricate with maximum tense quality. In clusters; not preceded by nasals. .
20	ഢ [dʰ]	[dʰ]	Metadata: Voiced aspirated sublamino-postalveolar stop. Medially in a small number of Sanskrit loans .	27	ജ [j]	[cʰ]	Metadata: Voiceless aspirated tense palatal affricate, occurs initially .
		[dʰ]	Metadata: In consonant clusters.			[J]	Metadata: Voiced tense palatal affricate in consonant clusters.
20	ഠ [dʰ]	[dʰ]	Metadata: Voiced aspirated sublamino-postalveolar stop. Medially in a small number of Sanskrit loans .	28	ഢ [jh]	[jh]	Metadata: Voiced palatal lax affricate. Initially before a vowel, and intervocally.
							Metadata: Voiced aspirated lamino-palatoalveolar affricate. Found only in Sanskrit loans Initially and medially

29	ഞ [ɲ]	[ɲ]	Metadata: Voiced lamino – palatal nasal. 1 Word initially 2. In the sequence -ɲj - in Sanskrit loans, where there is alternative of a long palatal nasal. 3. In the initial and medial sequence - ɲɲ 4. In the sequence - ɲc - in a small number of native words forms. 5. as a geminate consonant in native words. Palatal nasal contoid. 1 Word initially (short) 2. Short when followed by other palatal consonant 3. Long in intervocalic position.
30	ശ [ʃ]	[ʃ]	Metadata: Voiceless lamino palatal alveolar (retroflex) . Voiceless palatal sibilant slit fricative. Occurs initially and medially in loans. In initial position followed by a vowel in clusters, intervocalically. [long/short].
31	യ [y]	[y]	Metadata:- Voiced close front dorso – palatal semivowel . 1. occurs word initially: 2. Intervocalically 3. Medially as a geminated consonant 4. Finally as a variant of / j / 5. In word initial and medial clusters. Voiced palatal continuant . occurs short in initial and final position 2. short and long in medical position, in consonant clusters, and intervocalically position.
32	ക [k]	[k]	Metadata: Voiceless dorso velar plosive in initial position, and medially when doubled. Voiceless tense velar contoid; in initial position.
		[kj]	Metadata :Voiceless palatalised dorso velar plosive in the environment of preceding front vowel. Voiceless, most tense palatalised velar plosive contoid in the environment of preceding front vowel/in the sandhi environment where [j] precedes.
		[ɣ̥]	Metadata: Voiced dorso velar approximant in intervocalic position. Variant realisations in this environment are [h] and [ɦ]. Not palatalised, not following high front vowel.
		[g]	Metadata :Voiced dorso velar stop when preceded by a nasal. Not palatalised, velar stop contoid with a little voiced and lax quality in the environment preceded by a nasal.
		[t]	Metadata :The sequence [kʃ] [കഴ] is pronounced with retroflexion.
		[K]	Metadata :Voiced velar contoid with most tense quality, in clusters except the contoid occurs after [j] or a nasal.

33	ഖ [kh]	[kh]	Metadata: Voiceless aspirated tense velar plosive contoid. Intervocalically and when preceded by a nasal.
		[Kh]	Metadata: Voiceless aspirated tense velar plosive. Initially.
		[Kh]	Metadata: Voiceless aspirated, more tense velar plosive. Elsewhere .
34	ഗ [g]	[G]	-
		[g]	Metadata :Voiced and lax elsewhere ie, intervocalically and initially .
35	ഘ [gh]	[gh]	Metadata: In borrowing from Sanskrit, occurs in initial and medial position. VRP opines that it is represented in orthography only, not realized in speech. However, he considers it as an allophone of /Kh /]
36	ങ [ŋ]	[ŋ]	Metadata: Voiced dorso – velar nasal. 1. In the sequence - ŋg - in Sanskrit loans.2. In the Sequence - ŋk - bridging a morpheme juncture in a small number of word forms in the native lexicon. 3. As a geminated consonant in native words. 4. In the sequence - ŋk - in English loans.
		[ŋj]	Metadata: Voiced dorso – palato velar nasal. 1. In native lexicon, where geminate [ŋ] follows a front vowel.
		[ŋ<]	Metadata: Pre velaric nasal contoid with clear palatalization and tense quality. 1 In the names of fruits and plant except the one which occurs after a long low vowel.
		[ŋ>]	Metadata: Post velaric nasal contoid. 1 Long and tense when after a vowel in low-back region 2 Short before homorganic plosive.
37	ഈ [h]	[H]	Metadata :Voiceless extremely short velar fricative . – Finally.
		[h]	Metadata: Voiceless velaric or glottal fricative. 1. Initially 2. After vowel 3. In clusters 4. Intervocalically

IV. TEMU MALAYALAM PHONETIC ARCHIVE

In this work, an inclusive Malayalam phonetic data set which is being designed and developed as part of the Malayalam phonetic archive project owned by Thunchath Ezhuthachan Malayalam University (TEMU), Kerala, India is used for experimental purposes [9]. It is a fairly comprehensive database created by taking into consideration of a carefully compiled inventory of phones which are currently employed in

the Malayalam language. Malayalam phoneme segments are recorded in its standardized orthography followed by a number of examples of its occurrence in phonologically relevant different positions. Allophones are listed together and pronunciation of each example recorded from the natural speech is demonstrated in both male and female voices. The data comprises of 11 vowels, 2 diphthongs and 38 consonants, and its allophonic variation with 900 spoken words as examples. The following section describes the durational properties of the Malayalam vowel allophones derived on the basis of the detailed analysis conducted on the TEMU dataset. This archive is presently available in public domain under creative commons license. The dataset is archived and published in web portal [9].

V. DURATIONAL PROPERTIES OF MALAYALAM CONSONANT ALLOPHONES

Phoneme segmentation is the most important pre-processing in the phoneme level speech recognition. In phoneme segmentation algorithms, mostly the phonemes are assumed to be of the same length and segmented using a fixed size window. Durational analysis of phonemes performed in many languages reveals the variability in the duration of individual phonemes [12]. So the phone segmentation algorithms must consider the variability in phoneme duration for a better result. The phoneme level duration variability is language specific. Considering these facts a detailed analysis is conducted to establish durational phoneme models for the Malayalam language.

This section consolidates the detailed investigation performed on the duration of Malayalam consonant speech segments. The methodology of analysis is the same as employed for duration analysis of vowel allophones in the work by vivek *et.al.* [8] and is consolidated is Table IV.

TABLE IV. DURATIONAL STATISTICS OF MALAYALAM VOWEL ALLOPHONES

Sl. No.	Consonants	Allophone	Average Duration in Seconds	
			Male	Female
1	P ഹ	[p]	0.01340	0.03024
		[β]	0.02416	0.02316
		[b]	0.03509	0.03309
		[P]	0.03216	0.03016
2	p ^h ഹ	[p ^h]	0.05521	0.05321
3	b ങ	[B]	0.08418	0.08218
		[b]	0.07514	0.07714
4	b ^h ങ	[b ^h]	0.07674	0.09839
5	m മ	[m ^h]	0.07099	0.03903
		[M]	0.09796	0.06659
		[m]	0.09531	0.09075
		[m]	0.08339	0.08211
6	v ള	[w]	0.03924	0.03214
		[v]	0.08288	0.07236

7	t ത	[t]	0.01341	0.02404
		[t ^h]	0.03007	0.03109
		[ð]	0.02965	0.02846
		[d]	0.01431	0.02676
8	t ^h ത	[t ^h]	0.06104	0.06526
9	d ങ	[d]	0.04851	0.02513
		[d]	0.02243	0.02335
10	d ^h ങ	[d ^h]	0.04455	0.03448
11	n ന	[n]	0.08948	0.07447
		N	0.10796	0.15118
12	r റ	[d]	0.02843	0.02232
		[r]	0.01488	0.01821
13	n ന	[n]	0.16886	0.15322
		[n]	0.12224	0.14245
14	s സ	[s]	0.11396	0.13874
15	r റ r റ	[r]	0.10939	0.12834
16	l ല/ഛ	[l]	0.14006	0.12209
17	t ട	[d]	0.02822	0.01282
		[t]	0.01426	0.01430
		[t]	0.10374	0.10111
		[T]	0.02006	0.01256
18	th റ	[t]	0.02513	0.03250
		[T]	0.03658	0.03232
19	d ള	[d]	0.01608	0.02910
20	d ^h ള	[d ^h]	0.04557	0.05465
21	ṅ ണ	[ṅ]	0.08460	0.09675
22	ṣ ഷ	[s ^h]	0.14175	0.15000
23	l ള/ഛ	[l]	0.06460	0.04564
24	z ഴ	[z]	0.08493	0.09225
24	c ച	[c]	0.04773	0.05961
		[ç]	0.03856	0.04505
		[ç]	0.04609	0.05339
25	c ^h ച	[C ^h]	0.08531	0.06377
		[c ^h]	0.09809	0.08155
26	j ജ	[C ^h]	0.09645	0.08601
		[j]	0.06379	0.07052
27	jh ങ	[j]	0.05652	0.05273
		[j ^h]	0.08089	0.09455
28	j ^h ങ	[j ^h]	0.08089	0.09455
29	ṅ ണ	[ṅ]	0.09071	0.09066
30	ṣ ഷ	[ṣ]	0.11473	0.16944
30	y യ	[y]	0.06059	0.09967
31	k ക	[k]	0.02758	0.02120
		[kj]	0.04233	0.02900
		[ç]	0.05296	0.04231
		[g]	0.03121	0.02515

32	k ^h ക	[t]	0.02236	0.02452
		[K]	0.02247	0.02217
		[k ^h]	0.05651	0.05491
		[K ^h]	0.06391	0.05405
		[K ^h]	0.07139	0.05729
33	g ഗ	[G]	0.02534	0.05021
		[g]	0.05686	0.08251
34	g ^h ഘ	[g ^h]	0.07507	0.07607
35	η ണ	[η]	0.13133	0.13329
		[ηj]	0.10551	0.11969
		[η^v]	0.09054	0.15796
		[η>]	0.14165	0.15324
		[η]	0.15364	0.18484
36	h ഹ	[H]	0.0857	0.0854
		[h]	0.08746	0.08921

The first 4 phonemes in each vargga class (except nasals) and alveolar /t/ ട combine to form the plosive set in Malayalam. The nasals (/m /മ, /n /ന, /ɲ/ ണ/ണ/ണ, /ŋ/ ണ), fricatives (/s /സ, /ʃ/ ഷ, /ʒ/ ഷ, /h /ഹ), trills (/r/ര/ര), laterals (/l /ല/ല), approximants (/ɻ/ഴ) and Glides (/v/വ, /y/യ) are the remaining linguistic consonant classes in Malayalam. The durational analysis based on this classification is performed. The average duration of plosive is around 40 ms. Table V shows the average duration of each class of consonants. From the table it is evident that plosives have the smallest duration compared to other classes of consonants.

TABLE V. AVERAGE DURATION OF EACH CLASS OF CONSONANTS

Consonant Class	Average Duration in second (male)	Average Duration in second (female)
Plosive	0.03959	0.04004
Nasal	0.10894	0.11574
Fricative	0.10872	0.12656
Trill	0.10939	0.12834
Lateral	0.10233	0.08386
Approximant	0.08493	0.09225
Glide	0.060903	0.06805

The range of consonant duration varies from 13.3 ms to 168.4 ms for male and 12.4 ms to 184.3 ms for females. The average duration of consonants is much smaller compared to that vowels. A vargga classification also exists in Malayalam for consonant phonemes. There are 5 vargga classes in Malayalam. *Kavarggam* (/k/ ക, /k^h/ക, /g/ ഗ, /g^h/ഘ, /ŋ/ ണ), *chavarggam* (/c/ ച, /c^h/ച, /j/ ജ, /j^h/ഝ, /ɲ/ ണ), *tavarggam* (/t/ ട, /t^h/ട, /d/ ഡ, /d^h/ഢ, /ɳ/ ണ), *thavarggam* (/t/ ത, /t^h/ത, /d^h/ധ, /d/ദ, /ɳ/ ന) and *and pavarggam* (/p/ പ, /p^h/പ, /b/ബ, /b^h/ഭ, /m/ മ) each consisting of 5 consonants. The 5 consonants in each vargga are characterised linguistically as -voiceless unaspirated, -aspirated, +voiced unaspirated, +aspirated and Nasal (eg. k ക: -voiceless unaspirated, k^h ക: -aspirated, g ഗ: +voiced unaspirated, g^h ഘ: +aspirated, η ണ: Nasal). A similar durational pattern exists across voiceless unaspirated, -aspirated, +voiced unaspirated, +aspirated and nasal consonants in a vargga. Figure I shows the durational pattern of *ka-vargga* class.

From the experimental results, it is evident that the Malayalam consonant and vowel allophone duration information can be used for the development of different speech processing systems in Malayalam.

VI. CONCLUSION

Every phone in any spoken language is pronounced as one of its allophone. For the very reason, the properties of allophonic variations of each phone are very vital in continuous speech recognition and speech synthesis studies. In this work Malayalam vowel allophones are identified, classified and analyzed based on their durational properties. From the experimental results, it is evident that allophonic variations exist in the durational property of vowel phonemes. This aspect can be effectively used to improve the performance of the Malayalam ASR and synthesis. This work can be considered as a first step towards a paradigm shift to allophone based Malayalam speech processing.

A detailed investigation is performed on the duration of allophonic variations of Malayalam Consonant Phonemes in this paper. The durational properties obtained as part of the experiments can be effectively used for developing speech recognition and synthesis systems in Malayalam. Rule set for the formation of Malayalam consonant allophones are derived as part of this study. The process involved in the creation of TEMU dataset is also explained in detail. Durational properties of the Malayalam consonant allophones are analysed. The phoneme and allophone duration patterns emerged from the statistical analysis of speech corpora are presented. It is observed that the average duration of consonants is much

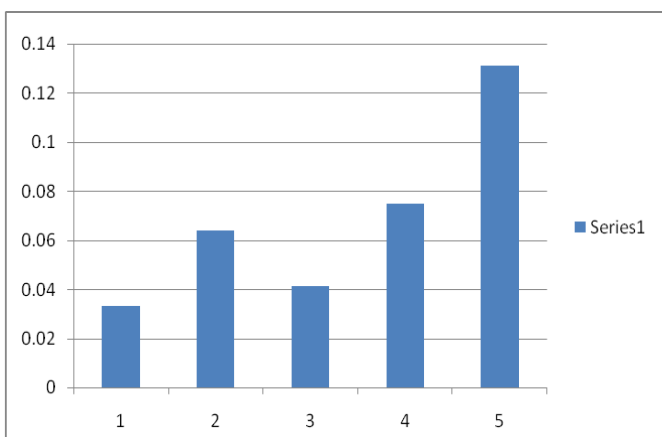


FIGURE I. THE DURATIONAL PATTERN AMONG CONSONANTS IN A VARGGA CLASS, 1. -VOICELESS UNASPIRATED, 2. -ASPIRATED, 3. +VOICED UNASPIRATED, 4. +ASPIRATED, 5. NASAL

smaller compared to those vowels. The range of consonant duration varies from 13.3 ms to 168.4 ms for male and 12.4 ms to 184.3 ms for females. Plosives have the smallest duration compared to other classes of consonants. The average duration of plosive is 40ms.

ACKNOWLEDGMENT

We would like to thank The Centre for Malayalam Language Technology, Thunjath Ezhuthachan Malayalam University, Kerala, India for providing necessary support to work with the TEMU Malayalam phonetic archive project and for granting permission to use same dataset for experimental purposes.

REFERENCES

- [1] "Report of the Commissioner for linguistic minorities: 50th report (July 2012 to June 2013)". Commissioner for Linguistic Minorities, Ministry of Minority Affairs, Government of India. Retrieved 17 September 2016.
- [2] M. Chu and Y. Feng, "Study on factors influencing durations of syllables in Mandarin," in Seventh European Conference on Speech Communication and Technology, 2001.
- [3] V. R. Prabodhachandran Nair, *Svanavijnjanam (Phonetics)*, SIL, 1980
- [4] D. H. Klatt, "Synthesis by rule of segmental durations in English sentences," *Front. Speech Comm. Res.*, pp. 287–299, 1979.
- [5] K. Bartkova and C. Sorin, "A model of segmental duration for speech synthesis in French," *Speech Commun.*, vol. 6, no. 3, pp. 245–260, 1987.
- [6] A. R. M. Simões, "Predicting sound segment duration in connected speech: An acoustical study of Brazilian Portuguese," in *The ESCA Workshop on Speech Synthesis*, 1991.
- [7] A. Bellur, K. B. Narayan, K. R. Krishnan, and H. A. Murthy, "Prosody modeling for syllable-based concatenative speech synthesis of Hindi and Tamil," in *2011 National Conference on Communications (NCC)*, 2011, pp. 1–5.
- [8] Vivek P, Sandesh E Pa, V L Lajish, "Durational Characteristics of Allophonic Variations in Malayalam Vowel Phonemes", 2018. 2348-2281.
- [9] "Malayalam Phonetic Archive," 2016. [Online]. Available: <http://www.cmltemu.in>.
- [10] E. Cosatto, G. Potamianos, and H. P. Graf, "Audio-visual unit selection for the synthesis of photo-realistic talking-heads," in *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No. 00TH8532)*, 2000, vol. 2, pp. 619–622.
- [11] Asher, Ronald E., and T. C. Kumari. *Malayalam*. Psychology Press, 1997.
- [12] Pols, Louis CW, Xue Wang, and Louis FM ten Bosch. "Modelling of phone duration (using the TIMIT database) and its potential benefit for ASR." *Speech Communication* 19.2 (1996): 161-176.