Acoustic Phonetic Characteristics of Kannada Language

Hemakumar G.

Assistant Professor, Department of Computer Science, Govt., College for Women, Mandya.

Abstract: Acoustic phonetics studies the physical properties of sounds and provides a language to distinguish one sound from another in quality and quantity. A study of acoustic characteristics of kannada begins with the phonemic analysis of the language. Phone to Phone relationship level is studied using Fifteen million words text corpus. Four hundred and forty five simple and most frequently using kannada words were selected and used to measure duration, intensity, frequency and formants of the phonemes in different emotional status and at different levels. Praat software and Pearl software were used for the purpose of analysis of output.

Key Index: Intensity, Frequency, peaks and Formants.

Introduction

Kannada language is one of the major Dravidian languages of India and it has 27th place in most spoken language in the world. There is clear distinction between the spoken and written forms of the language. Spoken Kannada tends to vary from region to region. The written form is more or less constant throughout Karnataka. However, the Ethnologue reports "about 20 dialects" of Kannada. The language uses forty nine phonemic letters, divided into three groups: *Swaragalu* (thirteen letters); *Yogavaahakagalu* (two letters); and *Vyanjanagalu* (thirty-four letters), similar to the vowels and consonants of English, respectively. The character set is almost identical to that of other Indian languages. The script itself, derived from Brahmi script, is fairly complicated like most other languages of India owing to the occurrence of various combinations of "half-letters" (glyphs), or symbols that attach to various letters in a manner similar to diacritical marks in the Roman languages. The Kannada script is an example for phonetic language, but for the sound of a "half n" (which becomes a half m). The number of written symbols, however, is far more than the forty-nine characters in the alphabet, because different characters can be combined to form *compound* characters *(ottaksharas)*. Each written symbol in the Kannada script corresponds with one syllable, as opposed to one phoneme in languages like English. The script of Kannada is also used in other languages such as Tulu, Kodava, Takk and Konkani.

I *Classification of Kannada Vowels (Swaragalu) and Consonants Speech Sounds:*

In the production of vowel sounds the air-current coming from lungs is allowed to go out without any obstruction in the mouth. However, there is a definite pattern in the production of different vowels sounds. Direction in which the tongue moves and variation in the shape of lips result in the change in the shape of the air chamber. It is this particular change which is responsible for the above mention production. In the production of Kannada vowels the vocal cords are vibrated and the nasal passage is closed. Classification of vowels is based on the position of tongue and lips. Table 1 shows the details of the classification of vowels and Table 4 shows the details classification of Consonants.

II Text Corpus and Speech Database creation:

Text corpus of 10 million words has collected from Dr. K. Naryana Murthy, Professor, Department of Computer and Information science, University of Hyderabad, Hyderabad, India. 5 million text corpuses has created by one month popular daily kannada newspapers (Prajavani and vijaya Karnataka), and few books available in Internet. All forty nine phones are independently uttered and four hundred and forty five simple and most commonly using words are selected to record at sampling rate of 8 KHz, 16 bps, mono channel. Each word was uttered by five male and five female of different age group. One female voice is recorded in different manner (angry, normal and sorrow).These signals were recorded at a little noisy environment, while Gold Wave Software was used to record with the help of mini microphone of frequency response 50 – 12500Hz.

III Acoustic Phonetic Recognition

In automatic speech recognition, the recognition of the word is primary. The word W is computed in a given observation sequence O by taking the product of two probabilities for each word, and choosing the word for which this product is greatest. These two terms have names the following names: P(W), is called the prior probability, which also the language model. P(O|W) is called the observation likelihood or acoustic model. The correlation between the value and recognized word is given by the formula.

$$\hat{W} = \arg\max_{W \varepsilon L} P(O|W)\, P(W)$$

To compute the acoustic model P(O|W), we need to make the assumption that the input sequence is a sequence of phones, rather than a sequence of acoustic observations. Observations are strings of phones, and produced the probability of these phone observations given a single word.

IV Interpretation of Sound Wave

Sound wave is interpreted with the help of spectrogram. It consists of Intensity, Frequency and Formants of sound wave. A brief description of these elements follows.

Let us begin with Spectrogram.

Spectral peaks are easily visible in a spectrum. They are distinct characteristics of different sounds. Phones have characteristic spectral "signatures". A spectrum shows the frequency components of wave at one point of time, whereas spectrogram is a way of envisioning how the different frequencies which make up a change in waveform over a period of time. Analysis of phonemes showed in table 2, 3, 4 and 5.

Several aspect of spectrogram are discussed below.

A. Measurement of Intensity:

Intensity is the acoustic power which is measured in watts. With reference to perception, they are computed in terms of decibels (dB). The simpler way of expressing the variation in intensity of vowels and consonants in various positions of their occurrences is through measuring the height of the highest period or peak from the central line, in millimeters (i.e. only the positive side).

B. Frequency calculation:

Frequency indicates the occurrence of the total number of complete cycles per second (cps). It is directly proportional to pitch. The frequency is calculated by the formula.

f = n / T, where f is frequency, n is the number of periods and T is time in seconds.

C. Formants:

A single energy maximum is called a formant (this definition is from Gunnar Fant). That is, for different sounds the maximum acoustic energy is distributed at different levels of frequency in different shapes and varying degrees of intensity. Both the shape and intensity, which are represented by the varying degrees of darkness, are important for the acoustic study. The shapes of the formants of vowels depend upon the following and preceding consonants. The harmonics are also present on the spectrograms, generally the harmonics run parallel to the fundamental pitch (F_0) or the frequency of glottal tone.

The formants are generally darker than the harmonics and they may have rising, level or falling shape, or the combination two or more shapes. Sometimes depending upon the phonetic environment, they may have a bend in the beginning. These characters make the formants distinct from the harmonics. It also requires a little practice to decipher the fused formants. Some part of a formant or the whole of it may fuse with or overlap on another formant. In such cases generally there would be a clue at the end of the formant. That is, the centre of the formant arrows out. The level at the arrow head could be taken as the centre of the formant.

The results analysis of average Formant frequencies of F0, F1, F2 and F3 of Vowels and Consonants are showed in table 4.

Table 1 Kannada vowels classification

Phonetic Classification	Vowels phone
High Vowels	/ ii / , / uu /
Higher-mid vowels	/ e / , / o /
Low vowel	/ aa /
Rounded vowels	/ u / , / o /
Unrounded vowels	/ i /, / e / , / a /
Dipthong	/ ai /, / au /
Additional Vowel	/ r /

V. Experimental Results:

The writing is based on the concept of *akshara* or the 'graphic syllable', which has a vowel as the final constituent, i.e. V, VCV, CV, CCV, CCCV etc. Word-initial V is written in its primary form; in the post consonantal position. The vowel is represented by a diacritic. The following observations relate to modern scripts of Kannada:

1. The vowels are written in their primary form when they occur at the beginning of a word. In all post consonantal positions only the secondary forms or diacritics are used.
2. The first member of a consonant cluster occurs in its primary form. All other consonants that follow it occur in their secondary (diacritic) form. The vowel which ends the orthographic syllable is added to the first member of the cluster.
3. The vowel diacritics are added to the consonant symbol to the top or right, but the secondary consonants are not allowed to touch the body of the consonantal symbol.
4. Vowels alone can become meaning word, but consonants alone cannot become words or akshara. A consonant needs help of vowels to become akshara.
5. The Liquid Trill phone /r/ is the only one phone which works with any phoneme combination.
6. In VC (vowel followed by consonant phone) combination the vowel phone /R/ and /au/ never combine with consonant phone /jh/ while forming meaningful kannada word.
7. In CV constituent the consonant phone /nG/ will never occurs with vowels /ee/, /o/, /oo/, /M/ and /H/. The consonant phones /nY/ and /N/ will never occur with vowels /R/, /M/, and /H/ and the consonant phone /L/ and /Sh/ will never occur with vowel /R/.

In the **Emotional case**:

The result of the study reveals a definite pattern in the relation when vowels are uttered duration and intensity varies in a uniform manner giving the emotional status of the subject. In the case of angry man duration and intensity of vowels is maximum, those of normal subject are found at middle level and those of sorrowful subject at lowest level. When angry men uttered consonants only intensity will be maximum, those of normal subject are found at middle level and those of sorrowful subject at lowest level of intensity and level of duration will almost varies, this details are showed in table 5.

Table 2: Comparison of kannada vowels with respect to duration (in milliseconds), intensity (in millimeters) and frequency (in cps); there relation with the position of their occurrence.

Sl. No	Phone's	Initial			Middle			Final		
		Duration	Intensity	Frequency	Duration	Intensity	Frequency	Duration	Intensity	Frequency
1	/ i /	75	6	125	60.77	9	120	80.81	7	125
2	/ ii /	132	7	121.21	136.41	8.5	137.68	138.16	9	114.28
3	/ e /	114	8.5	118.18	83.16	8.5	137.5	118.85	8.5	117.64
4	/ ee /				151.16	9.5	135.48			
5	/ ai /	75.14	5	112.5	64.08	7.5	107.69			
6	/ u /	64.73	5.5	127.27	58.05	7	136.2	84.98	3	94.12
7	/ uu /	150	6	123.66	168	8	142.85			
8	/ o /	98	9	132.65	84	8.5	106.25	138.06	8.5	137.68
9	/ oo /	196.66	7	115.78	146.22	8	127.9			
10	/ a /	67.13	8.5	142.85	71.84	8	128.56	68.54	6	100
11	/ aa /	169.5	8.5	117.64	157.8	8.5	106.25	138.06	8.5	137.68

Table 3: Comparison of kannada Consonants Duration (in seconds), intensity (Int: in millimeters) and frequency (in cps) are related to the position of their occurrence.

Phoneme	VCV			C followed by C			C1C1(consonant + vattaksharas)		
	Duration	**Int**	**Frequency**	**Duration**	**Int**	**Frequency**	**Duration**	**Int**	**Frequency**
/p/	107.50			109.75			198		
/b/	62.50			107.11			152.8		
/t/	106.50			120.16			165.5		
/d/	56.43			105			147		
/k/	107.25			129.66			190.57		
/g/	56.78			70.33			132		
/T/	82.00			104.77			169.4		
/D/	33.90			50.37			135		
/c/	51.66			49			65	5.5	
/j/	31.25			32	3		37	3	
/h/	44.20	2							
/s/	101.33	2.5					150	3	
/sh/	142.33	5.5					153	6	
/Sh/	98.00	5							
/m/	54.66	6	133.33	91.77	3.5	116.66	117.5	6	144.06
/n/	46.80	5.5	150	88.88	5	130.71	128	6	132.81
/N/	55.80	8	125.22	85.9	4	142.85	89.33	5	123.07
/l/	53.90	6	118.47	84.66	4.5	111.11	15	6	106.66
/L/	48.33	8	109.1				63	7.5	123.07
/r/	20.30	1	133.33						
/v/	60.07	6	111.11				75	6	123.52
/y/	65.44	8	116.66	63.6	7.5	142.85	120	7	180

Table 4: Formant frequencies of F0, F1, F2, and F3 of Vowels and Consonants of Kannada

Type	Phone's	Formants in Hertz			
		F0	F1	F2	F3
Vowels	a	866	1347	1708	2973
	aa	871	1319	1808	3072
	i	387	1622	2620	3183
	ii	412	1657	2665	3287
	u	523	1150	1881	3050
	uu	405	997	1869	3067
	R	509	1261	1909	2993
	e	556	1533	2312	3036
	ee	572	1535	2374	3036

Dipthongs		ai	690	1733	2168	3117
		au	605	1182	1753	2927
Plosive	Unaspirated	p	784	1321	1668	2969
		b	788	1324	1697	2932
		t	787	1364	1760	3048
		d	705	1264	1715	2962
		k	803	1320	1721	2956
		g	764	1308	1674	2946
		T	792	1351	1770	2846
		D	755	1386	1778	3034
	Aspirated	ph	766	1209	1758	2981
		bh	789	1375	1678	2916
		th	828	1181	1784	3153
		dh	746	1324	1789	3002
		kh	810	1263	1832	3063
		gh	671	1258	1787	2929
		Th	864	1282	1869	3082
		Dh	830	1295	1798	2994
Affricates	**Unaspirated**	c	727	1297	1804	2825
		j	744	1371	1774	2933
	aspirated	ch	835	1257	1757	3026
		jh	707	1437	1792	2877
Fricattives		h	828	1144	1817	2988
		s	784	1310	1731	3011
		sh	710	1330	1815	3105
		Sh	777	1449	1853	3131
Nasals		m	864	1202	1808	3208
		n	850	1387	1756	3203
		N	816	1274	1732	3057
		nY	773	1576	1857	3030
		nG	778	1285	1833	3122
Liquids		l	806	1466	1787	2904
		L	605	1501	1831	2978
		r	762	1390	1783	3011
Semi Vowels		v	784	1192	1798	2925
		y	733	1456	1812	2989

Table 5: Comparison of kannada Consonants Duration and intensity occurrence Independently during Normal and Sorrow condition.

Type			uttered at Normal time		uttered at Sorrow time	
		Phone	**Dur**	**Int**	**Dur**	**Int**
Plosive	Unaspirated	p	0.149125	72.99	0.226375	64.21
		b	0.189125	73.46	0.190375	58.19
		t	0.233	70.79	0.206625	52.75
		d	0.213	64.26	0.137125	54.13
		k	0.140125	69.29	0.161125	57.79
		g	0.17175	66.2	0.233	60.79
		T	0.141	71.84	0.275625	56.38
		D	0.29825	63.6	0.161125	58.23
	Aspirated	ph	0.31825	73.58	0.284875	57.67
		bh	0.21575	68.61	0.28225	59.2
		th	0.35025	70.86	0.271625	57.16
		dh	0.283625	69.87	0.34625	57.88
		kh	0.27425	61.66	0.352875	59.5
		gh	0.251625	66.67	0.284875	61.44
		Th	0.324875	70.93	0.322125	59.64
		Dh	0.29825	63.6	0.269	56.62
Affricates	**Unaspirated**	c	0.24625	65.97	0.13325	54.71
		j	0.238375	65.56	0.17575	54.15
	aspirated	ch	0.275625	74.22	0.25975	59.51
		jh	0.197125	65.01	0.245	59.09
Fricatives		h	0.237	64.7	0.23725	62.35
		s	0.21175	62.37	0.1785	56.4
		sh	0.303125	68.66	0.157125	53.76
		Sh	0.37675	69.51	0.221	62.18
Nasals		m	0.197125	68.63	0.249	53.58
		n	0.233	65.1	0.1665	55.89
		N	0.221	65.98	0.295625	58.27
		nY	0.19575	63.65	0.14775	52.33
		nG	0.242375	65.4	0.210375	60.54
Liquids		l	0.237	66	0.202375	61.89
		L	0.258375	67.56	0.26225	58.57
		r	0.249	69.11	0.241	54.22
Semi Vowels		v	0.265	70.41	0.169125	57.1
		y	0.217125	64.99	0.253	55.34

V. *Discussion and Conclusion:*

Vowels are classified into short vowels and long vowels. The long vowels showed 180-230 percent duration when compared with their short counterparts. The short vowels showed 50-70% more duration in the final position than their initial position when the vowel in the penultimate syllable was short. This was true even when the penultimate vowel was extra short due to the tendency of loss or when there was a total loss. When the vowel in the penultimate syllable was long the duration of final vowel was almost the same as in its initial position.

In spectrographic analysis all the sounds are interpreted on the basis of the formant structures of vowels only. The oscillographic analysis shows the duration, intensity and frequency of the sounds. But spectrographic analysis, gives a better insight about the nature of sounds. The sound spectrograph is a good apparatus to study the distribution of acoustic energy over a frequency range from 100 to 10,000 cps. Thus it is possible to determine the quality of vowels accurately.

VI. Acknowledgement

The author would like to thanks the Prof K. Narayan Murthy, Prof., Nandan, Prof., Punitha Swamy, friends and anonymous reviewers for their constructive comments and extremely helpful suggestions and the Editorial Staff for their efforts in preparation of the proof for this paper.

References

1. Lawrence Rabiner, Biing-Hwang Jung, "**Fundamentals of Speech Recognition",** PEARSON EDUCATION (Singapore) PRIVATE LIMITED, Indian Branch, 482 F.I.E Patpargans, Delhi 110092, India, © 1993.

2. Daniel Jurafsky and James H. Martin, **"An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition"**, PEARSON EDUCATION (Singapore) PRIVATE LIMITED, Indian Branch, 482 F.I.E Patpargans, Delhi 110092, India, © 2003.

3. T. Balasubramanian, "**A Textbook of English Phonetics for Indian Students",** Published by Rajiv Beri for MACMILLAN INDIA Ltd., 2/10 Ansari Road, Daryaganj, New Delhi 110002, First published 1981 (Reprinted 15 times).

4. U.P. Upadhyaya, **"Kannada Phonetic Reader"**, CENTRAL INSTITUTE OF INDIAN LANGUAGES, Mysore, Publication No. 467.

5. Patricia Scanlon, Daniel P.W. Ellis, "**Using Broad Phonetic Group Experts for Improved Speech Recoginition**", IEEE transaction on Audio, Speech and Language processing, VOL 15, No. 3, March 2007.

6. B.B. Rajapurohit, "Acoustic Characteristics of Kannada", CENTRAL INSTITUTE OF INDIAN LANGUAGES, Mysore, occasional monographs series No. 27.

7. Praat software and praat help manual.

8. D.J.Ravi and Sudarshan Patilkulkarni, "A Novel Approach to Develop Speech Database for Kannada Text-to-Speech System" , published by Int. J. on Recent Trends in Engineering & Technology, Vol. 05, No. 01, Mar 2011

9. M.A.Anusuya1, and S.K.Katti, "Mel Frequency Discrete Wavelet Coefficients for Kannada Speech Recognition using PCA", published by Proc. of Int. Conf. on Advances in Computer Science 2010.

10. "Statistical Analyses of Telugu Text Corpora" by *G:BharadwajaKumar*, *KaviNarayanaMurthy*, *BBChaudhuri*.

11. "A Computational Phonetic Model for Indian Language Scripts", by
Anil Kumar Singh, Language Technologies Research Centre IIIT, Hyderabad, India.

12. "Salient phonetic features of Indian languages for Speech Technology", by
PERI BHASKARARAO, Tokyo University of Foreign Studies, Tokyo, Japan.

13.http://brahmi.sourceforge.net/docs/KannadaComputing.html

14. http://en.wikipedia.org/wiki/Kannada

15.http://www.omniglot.com/writing/kannada.htm

16.http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers.

17.http://www.censusindia.gov.in/Census_Data_2001

18. Bhadriraju Krishnamurti, **"The Dravidian Languages"**, Published in the United States by Cambridge University Press, New York

Author Biographies : I Hemakumar G. B.Sc, M.Sc, M.Phil. I started my teaching profession from University Evening College, University of Mysore, Mysore and having 7 years' teaching experience and 3 years of research experience. At present I am working in the prestigious institution i.e., Government College for Women, Mandya. I am working for UGC Sponsored minor research project and My interest areas of research are Automatic Speech Recognition, Pattern Recognition and Digital Signal Processing.