Journal Of Harmonized Research (JOHR)

Journal Of Harmonized Research in Engineering
3(3), 2015, 116-121

ISSN 2347 – 7393

Original Research Article

A STUDY ON ESTIMATION OF FORMANTS OF ASSAMESE VOWEL AND WORDS USING LPC METHOD

Mousmita Devi[1], Prof. P.H.Talukdar[2]

Research Scholar, Department of Instrumentation & USIC, Gauhati University, Guwahati, India [1]
Professor, Department of Instrumentation & USIC, Gauhati University, Guwahati, India [2]

Abstract: In human speech one of the key aspects is its Formant structure. Formants are resonances of the vocal tract, and as such they have a close relation to the vocal tract geometry. In other words Formant frequency may be refers to the spectral peak of the sound spectrum. Generally, for perception and discrimination three formants namely First (F1), Second (F2) and Third (F3) are considered. In this paper, the formant frequency of Assamese vowels and words of typical structure i.e. CV, CVC, VCV are estimated which can be helpful for developing Assamese Automatic Speech recognition (ASR) system. In this study, we design a database which consist of eight Assamese vowels and words of 5 from each structure total 15 words uttered by 5 times which make our database having total 115 samples. In this study we have analyze the sample for F1, F2 and F3 formant frequencies. It is observed that a significant variation is present in the formant frequencies with respect to gender for both phoneme set as well as the word set.

Keyword: *Speech recognition, Formant frequency, word, phoneme, resonance, LPC.*

1. Introduction*:*

The Assamese or Asamyia (IPA:[x mija]) is a major language in the North-Eastern part of India with its own unique identity. Assamese language has some unique phonetic properties in their phonemes. There is other phonological uniqueness of Assamese pronunciation which shows minor variations when spoken by people from different regions of the state. This makes Assamese speech unique and hence requires an extensive study to develop a speech recognition and synthesis system in Assamese [2]. The ASSAMESE phonemic inventory consists of eight vowels, ten diphthongs, twenty-one consonants and two semi vowels **[**8**]**. Vowels are classified as front, mid, or back, corresponding to the position of the tongue hump [1]. The ASSAMESE vowels are shown in the table below.

For Correspondence:
devimausmita@gmail.com
Received on: September 2015
Accepted after revision: September 2015
Downloaded from: www.johronline.com

Table1. Vowels in Assamese language

Tongue Position → / Shape of Lips →		Front		Central		Back	
		Unrounded		Neutral		Rounded	
Height of the Tongue ↓	Space in the Oral Cavity ↓	IPA	Assamese Vowel Phoneme	IPA	Assamese Vowel Phoneme	IPA	Assamese Vowel Phoneme
High	Close	/i/	ই			/u/	উ
High-Mid	Half Close	/e/	এ’			/o/	ও
Low- Mid	Half Open	/ɛ/	এ			/ɔ/	অ’
Low	Open			/a/	আ	/ ɒ/	অ

Selection of Word List

In this study we have collected the speech samples which are phonetically rich as well as which are frequently used in our day today life. Total fifteen phonetically rich words have been selected five each in the three words structure i.e. CV,CVC and VCV respectively.

Assamese Word List considered to build in the present data sample is listed in the given tables.

Table2. Word List of CV type

WORD LIST	
CV	
Word	Meaning
[ma]	Mother
[kha]	Eat
[nao]	Boat
[nɔ]	Nine
[ga]	Body

Table3. Word List of CVC type

WORD LIST	
CVC	
Word	Meaning
[nas]	Dance
[gan]	Song
[nak]	Nose
[gɔs]	Tree
[dhɔn]	Money

Table4. Word List of VC type

WORD LIST	
VC	
Word	Meaning
[am]	Mangoes
[ah]	Come
[ak]	Draw
[ek]	One
[aji]	Today

The purpose of this study was to examine the formants of vowels and words of Assamese language. And also this study includes an acoustic comparative analysis of formant change during different position of vowel when comes under a word. The objective of this paper is to analyze the effect of vowel in a wordpresent in the Assamese language, by applying Formant frequency measure.

2.Methodology

Selection of sample

In the present study, the vowels and word sounds of the Assamese language for both female and male has been collected for analysis. All speakers are within the age group of 20-25 years. The formant frequency depends upon the dimensions and shape of the vocal tract, where each shape is characterized by a set of formant frequencies. For this reason we have selected the subjects are within the age group of 20-25 years.The correct pronunciation is examined by some Assamese Phonetic experts.

Recording Procedure

The recording was done in a quiet room with a noise cancelling microphone using the recording facilities of a typical multimedia computer system. The boundaries of syllable were marked according to the pronunciation of the team. The voice was recorded and processed on Audacity and Cool Edit Pro. The equipment used in recording was a microphone with a frequency response of 48000 Hz with a 16 bit sound card and high quality speakers.Subjects were positioned approximately 8 to 10 inches from the microphone while recording.

Formant Estimation using Linear Predictive Coding(LPC) Method

In this study, we have studied the results for the formant frequencies of eight Assamese vowel phonemes as well as the different word structure

estimated from the 10th orderLinear Predictive Coding(LPC) spectrum envelop. The first three consecutive peaks are generally known as the first three formants. Theformant model used in the present study for the determination of Formant Frequency of Assamese vowels and words is based on themodel proposed by Welling et. al. [8]. Applying this technique, the entire frequency range is divided into a fixednumber of segments, where each of these segments represents frequency. A second order resonator for each segmentsK, with a specific boundary is defined. A predictor polynomial defined as a Fourier Transform of the Correspondingsecond order predictor is given by, [9]

$$A_k(e^{jw}) = 1-\alpha_k e^{jw} - \beta_k e^{-j2w} \quad \ldots(1)$$

Where, α_k and β_k are the real valued prediction co-efficient. The formant frequency is given by,

$$P_f = \text{across } [-\alpha_k(1 - \beta_k)/4\,\beta_k \quad \ldots(2)$$

The value of α_k and β_k are defined as,

$$\alpha_k<2 \quad \ldots (3)$$

And,

$$-1<\beta_k<[-|\alpha_k|/(4-|\alpha_k|)] \quad \ldots (4)$$

Now, using the above equations, the Formant Frequencies of the selected dataset are estimated for both Male and Femaleinformants. For obtaining the Formantfrequency, the spectra is subjected to First Fourier Transformation(FFT).

Table5. Formant frequencies for eight Assamese vowel phonemes corresponding to Assamese male and female informants.

vowel →		/ɒ/	/a/	/ɔ/	/i/	/ɛ/	/e/	/o/	/u/
Female	F1	362.921	226.651	170.651	348.167	234.9141	213.871	171.9985	159.4934
	F2	682.355	836.894	1676.894	1434.975	716.467	1479.67	526.397	804.855
	F3	1450.376	1639.841	2076.894	2239.841	1529.596	2361.139	1656.427	1264.095
Male	F1	301.3358	554.286	248.286	298.145	174.988	187.155	181.4476	146.5354
	F2	560.208	753.074	543.074	1078.836	786.854	1018.253	479.185	521.31
	F3	1246.696	1454.648	1154.648	1580.886	1310.888	1321.403	1387.472	1323.541

Table6.Formant frequencies for Assamese words (CV) corresponding to Assamese male and female informant

CV →		/ ma /	/kha /	/nao /	/nɔ /	/ga /
Female	F1	217.921	238.0311	102.651	328.167	202.9141
	F2	582.355	699.944	536.894	804.975	616.4673
	F3	1150.376	1125.062	939.841	1429.596	1270.473
Male	F1	101.921	118.0311	112.651	118.167	108.9141
	F2	472.363	458.944	329.786	689.995	445.454
	F3	943.342	1012.112	721.934	1289.126	1070.245

Table7. Formant frequencies for Assamese words (CVC) corresponding to Assamese male and female informants.

CVC →		/nas /	/gan /	/nak /	/gɔs /	/dhɔn /
Female	F1	148.453	137.5623	195.60	223.7294	219.126
	F2	511.654	423.543	667.320	889.611	798.677
	F3	956.778	976.852	1176.725	1492.121	1521.890
Male	F1	127.654	123.540	121.621	211.212	209.878
	F2	498.652	412.765	556.778	572.890	678.889
	F3	845.167	989.709	1190.765	1203.554	1349.890

Table8.Formant frequencies for Assamese words (VC) corresponding to Assamese male and female informants.

VC →		/ am /	/ ah /	/ak /	/ek /	/aji /
Female	F1	198.936	237.679	195.606	189.729	211.226
	F2	633.692	599.187	988.555	873.032	853.724
	F3	1668.225	1526.852	1425.725	1833.765	1014.383
Male	F1	177.966	256.123	182.102	209.808	301.936
	F2	418.731	542.706	879.948	672.465	700.978
	F3	1595.515	1622.574	1440.543	1638.462	1049.505

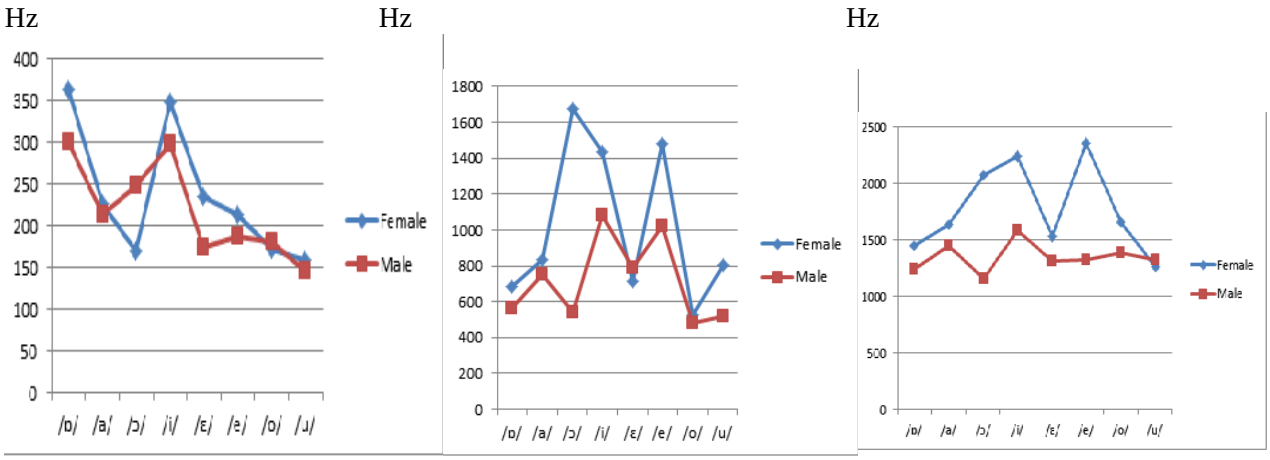

Fig.1.Variation of F1, F2, F3 with different vowels for male and female informants

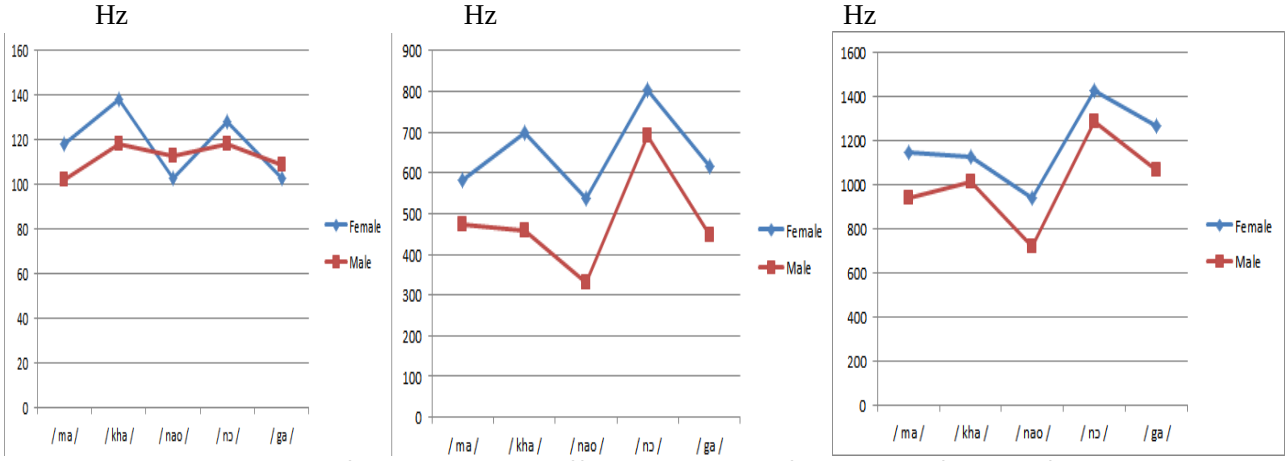

Fig.2.Variation of F1, F2, F3 with different words(CV) for male and female informants

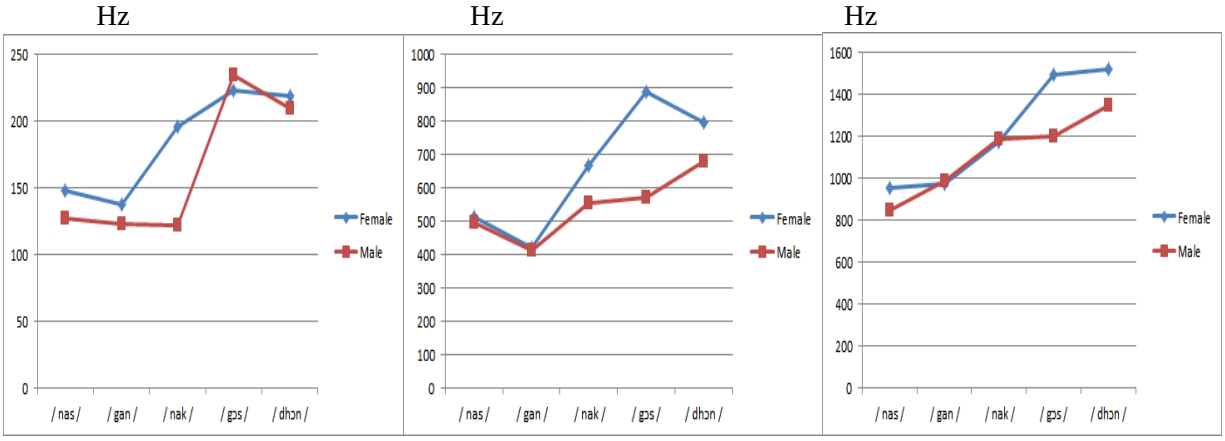

Fig.3.Variation of F1, F2, F3 with different words (CVC) for male and female informants

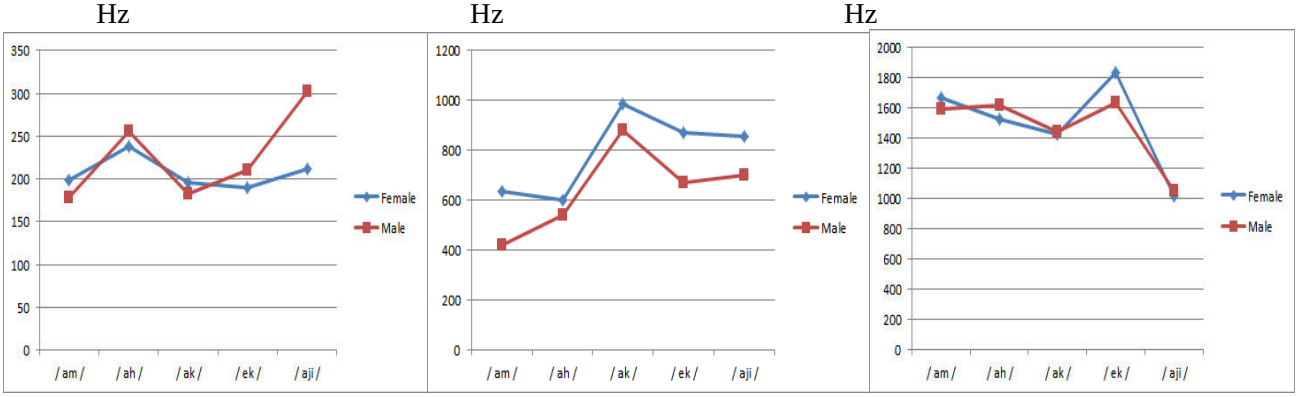

Fig.4. Variation of F1, F2, F3 with different words (VC) for male and female informants

4.Result and Discussion

Depending on the analysis on formantfrequencies of Assamesevowel phonemes and words thefollowing observations were made-

a) The formants are the resonance peaks in the corresponding LPC spectra. The frequencies of the first three formants contain sufficient information for the recognition of vowels as well as other word structures. For each of the vowel phoneme and word speech of both male and female informants, we extract the first three formants by locating the first three consecutive peaks in spectrally smoothed log spectra.

b) Table [1-4] depicts the values of the formant frequency for eight Assamese vowels and different word structure corresponding to Assamese male and female informants.

c) We have seen that the first three formants are distinctly different for all the eight Assamese vowel phonemes. Since different vowels have their formants at characteristic places, the spectrum can distinguish vowels from each other. The first formant F1 is associated with changes in mouth opening. Sounds requiring small mouth openings have low-frequency first formant (F1) and those requiring a wide mouth opening have high frequency F1s. For instance, the vowel /e/ requires a small mouth opening in contrast to the vowel /ɒ /. The F1 of the vowel /e/ is found to be about 213.871 Hz for female and 187.155 Hz for male informants, whereas the F1 of the vowel /ɒ/ is about 362.921 Hz for female and 301.3358 Hz for male informants. So as when wide opening vowel like /ɒ/ comes in constructing a word in most of the time its first formant found at higher frequency as compared to words have small mouth opening vowel.

d) The location of the first two formants (F1 and F2) plays a significant role in determining vowel identity, although the formants still differ from speaker to speaker.

We have also seen that the location of the first three formants is very different for male and female informants. Females have higher formant frequencies than males for all the vowels.

e) The investigation have shown that the range of formant frequency is maximum in case of isolated vowels, but when the vowels are placed in the nucleus of a structurelike CV, VC or CVC, the formant frequency decreases.

REFERENCES-

1. Rabinar L.R., Juang B.H., 'Fundamental of Speech Recognition', Dorling Kinderseley (India).W.-K. Chen, Linear Networks and Systems(Book style). Belmont, CA: Wadsworth, pp. 123–135, 1998.
2. D. Talkin, 'Speech Formant Frequency estimation using dynamic programming with modulated transition cost', AT&T Bell labs,McGraw Hill, NJ, 1987
3. Talukdar P.H; 2010. **Speech production, Analysis andCoding,** Lambert Publication, Germany 2010.
4. O. Schmidbaner, 'An algorithm for automatic formant extraction in continuous speech', Proc. EUSIPCO-90, Fifth European SignalProcessing Conference,pp-115, 1990.
5. Willing I., and Ney, II. **Formant Estimation for Speech Recognition**, IEEE Transactions on Speech andAudio Processing, Vol 6. pp.-36-48,1998
6. Atal, B. S. and Hanauer, S. L., 'Speech Analysis and Synthesis by Linear Prediction of the Speech Wave', J. Acoust. Soc. Am., 50, pp.637-655, 1971.
7. Snell R.C. and Milinazzo F., 'Formant Location form LPC Analysis Data', IEEE trans. Speech Audion, Processing, pp-129, 1993.
8. BanikantaKakati. Assamese,its formation and development, 3rd Ed 1972