



Analytical Study of Bodo Vowels – A Core Language of NE India Using Formant Frequency

Uzzal Sharma¹

Assistant professor, Dept. of CSE & IT, DBCET, Assam Don Bosco University, Guwahati, Assam, India¹

ABSTRACT: During the last few decades, a number of approaches (Formant Tracking, Articulatory model, and Auditory model) have been developed for the analysis of speech signal with a view to strategic development of speaker identification. Out of the above mentioned approaches, Formant Tracking Model has been able to attract more attention. Formant Frequency, based on Formant Tracking Model can be defined as the spectral peak of the sound spectrum $|P(f)|$. For the current study, the recorded data set is analyzed for First(F1), Second(F2) and Third(F3) formant frequencies. The observation reveals remarkable significance in F1 and F2 with respect to speaker and gender which in turn can be utilized in speaker and gender identification.

Keywords: Articulatory model, auditory model, Formant Tracking Model, FFT, Formant Frequency, Resonance Frequency.

I. INTRODUCTION

The human vocal tract can be considered as a single tube or combination of some tubes where the varying cross-sectional area is excited either at some end or at a particular point along the tube. According to the Acoustic theory, the transfer function of energy from the excitation source to the output can be referred as natural frequencies of resonances of the tube. The formant frequency depends upon the dimensions and shape of the vocal tract, where each shape is characterized by a set of formant frequencies. The variation in the shape of the vocal tract produces different sounds. Thus as the shape of the vocal tract changes, the spectral characteristics of the speech signal vary with time. Typically, a human vocal tract exhibits about three significant resonances below 3500 Hz. The formant frequency representation is a highly efficient and compact representation of speech sound [1].

Vowels of any language are very important towards the identification of speaker as it contains large amount of speaker information. In this paper, an attempt is made to analyze the vowels present in the Bodo language, a major language of NE India, by applying Formant frequency measure and some distinctive results were observed [2].

II. SPEECH RECOGNITION AND FORMANT ANALYSIS

Formants can be defined as the spectral peak of the sound spectrum $|P(f)|$ [5]. These are normally the peaks, which are known as the Resonance Frequency, $|T(f)|$. Although in most of the cases, it is seen that the Resonance Frequency, $|T(f)|$ and Formant Frequency, $|P(f)|$ is same, but in some particular cases it may be different.

During the last few decades, a number of approaches have been developed for analysis and synthesis of speech signal with a view to development of speaker identification system. Among all, Formant Tracking Method [3, 4], Articulatory model [6], and Auditory model are considered as the basic models for speech recognition and research. Among all, Formant Tracking Model based on Linear Predictive Coding (LPC) has been found to be more successful [5, 7]. The formant model used in the present study for the determination of Formant Frequency of Bodo vowels is based on the model proposed by Welling et. al. [8]. Applying this technique, the entire frequency range is divided into a fixed number of segments, where each of these segments represents frequency. A second order resonator for each segments K , with a specific boundary is defined. A predictor polynomial defined as a Fourier Transform of the Corresponding second order predictor is given by, [9]



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2013

$$A_k(e^{j\omega}) = 1 - \alpha_k e^{j\omega} - \beta_k e^{-j2\omega} \quad (1)$$

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

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

The value of α_k and β_k are defined as,

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

and,

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

Now, using equation (2), the Formant Frequencies of six Bodo vowels are estimated for both Male and Female informants. The correct pronunciation is examined by some Bodo Phonetic experts. For obtaining the Formant frequency, the spectra is subjected to First Fourier Transformation (FFT).

III. METHODOLOGY

In the present research, the content materials are prepared in two sets printed form – one for the informants and the other for the verifier. The informants were asked for to speak in correct stress-free pronunciation, while maintaining the constant pitch as far as possible. While preparing the content materials, the vowel sounds under investigation are embedded in the natural running words, than vowels are separated from the words and stored as a corpora entry. The informants are given a rest of 10 to 15 minutes after every session of recording containing a particular item.

The Male and Female informants of age between 15 to 30 years possessing a pleasant and good voice quality are chosen to record the data. Only native speakers being graduate or post-graduate are selected. To ensure accuracy and consistency, the recording process is supervised by acoustic phonetic experts of the language concerned.

IV. RESULTS

For the current study, the recorded data set is analysed for First(F1), Second(F2) and Third(F3) formant frequencies as under :

The recording and separation of the important segment is performed using audio editing software Cool Edit Pro and after that the analysis was done using MATLAB 7.1, and COLEA (subset of a COchLEA Toolbox), a special speech signal analysis tool belongs to MATLAB. Each digitized voice recorded, is divided into 50 frames of duration 20 millisecond (ms) each. Every frame contains approximately 441 samples and for each frame Formant Frequencies (F1, F2, and F3) are calculated and investigated. The variation of the formant frequencies for the Bodo vowels corresponding to the selected speakers have been shown in Table-I for male and Table-II for female and depicted in Fig. 1(a) and Fig. 1(b) for vowel /a/ and /e/ respectively when uttered by male informants and Fig. 2(a) and Fig. 2(b) for vowel /a/ and /e/ respectively, when uttered by female informants.

TABLE I
RANGE OF VARIATION OF FORMANT FREQUENCIES OF BODO VOWELS /A/ AND /E/ (MALE)

Vowel	Value	F1(KHz)	F2(KHz)	F3(KHz)
/a/	Max	1.5050	3.886768	3.981762
	Min	0.4681	1.129727	2.138752
	Average	0.9622	1.741653	3.499333
	Range	1.0368	2.757041	1.84301
/e/	Max	0.4665	2.451996	3.993117
	Min	0.2205	2.160764	2.621804
	Average	0.3897	2.274905	3.51991
	Range	0.246	0.291232	1.371313

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2013

TABLE II
RANGE OF VARIATION OF FORMANT FREQUENCIES OF BODO VOWELS /a/ AND /e/ (FEMALE)

Vowel	Value	F1(KHz)	F2(KHz)	F3(KHz)
/a/	Max	1.7042	3.8288	3.9979
	Min	0.1798	1.4184	2.8619
	Average	1.0301	1.7679	3.5056
	Range	1.5243	2.4105	1.1361
/e/	Max	0.0012	0.0028	0.0039
	Min	0.0003	0.0025	0.0034
	Average	0.0005	0.0026	0.0037
	Range	0.0009	0.0003	0.0005

V. DISCUSSION AND CONCLUDING REMARKS

The Formant Frequency analysis of Bodo Vowels with respect to male and female informants, reveals the following facts :

From the pictorial representation of Formant Frequencies of vowels for both Male and Female informants, it is observed that the Formant Frequency F3, do not carry any remarkable characteristic, for distinct Speech and individual Speaker. So, F3 does not play any significant role as far as Speech and Speaker Recognition is concerned. At the same time, the variation of F1, with respect to different vowels is quite distinct and prominent and upto some extent the F2 also shows some remarkable significance. Thus F1 and F2 can be considered as useful for speech and speaker identification of Bodo Male and Female informants.

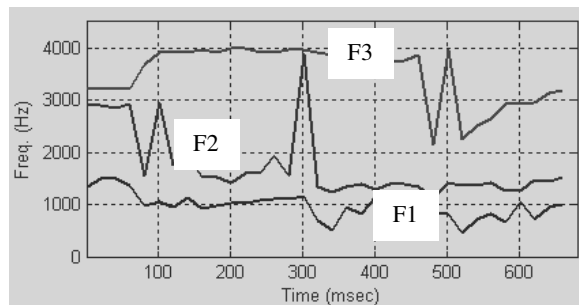


Fig. 1(a) Formant Frequencies of Bodo vowel /a/ (Male)

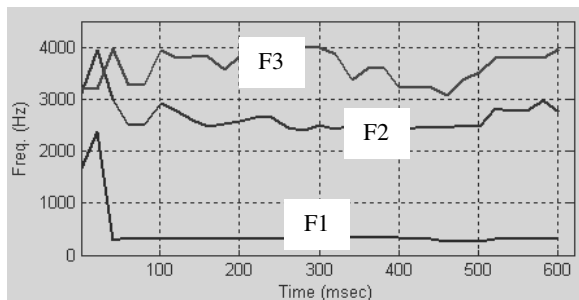


Fig. 1(b) Formant Frequencies of Bodo vowel /e/ (Male)

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 12, December 2013

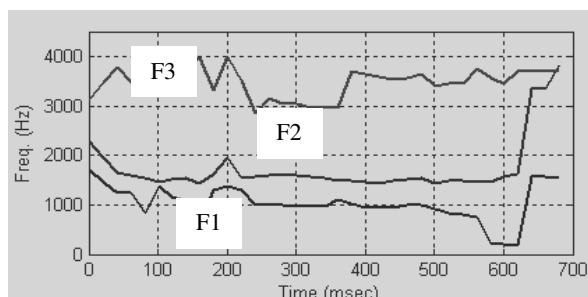


Fig. 2(a) Formant Frequencies of Bodo vowel /a/ (Female)

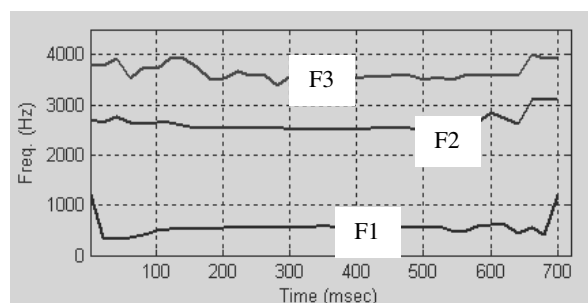


Fig. 2(b) Formant Frequencies of Bodo vowel /e/ (Female)

It is also observed that that second formant, F2 of /a/, /e/, /o/, /w/ plays important role in the gender identification of Bodo informants. On the other hand, in case of male informants, the change in the frequency is more gradual with one or two extinctions, and uniform frequency change is observed in all F1, F2, and F3.

REFERENCES

- [1] [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] [2] D. Talkin, 'Speech Formant Frequency estimation using dynamic programming with modulated transition cost', AT&T Bell labs, McGraw Hill, NJ, 1987
- [3] [3] O. Schmidbaner, 'An algorithm for automatic formant extraction in continuous speech', Proc. EUSIPCO-90, Fifth European Signal Processing Conference, pp-115, 1990.
- [4] [4] 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.
- [5] [5] H.B. Richard, Mason J.S. Hunt M. J. and Bridle J.S., 'Deriving Articulatory Representation of Speech', Proc. of European Conference of Speech Communication and Technology, Madrid, Spain, pp- 761, 1995.
- [6] [6] Snell R.C. and Milinazzo F., 'Formant Location form LPC Analysis Data', IEEE trans. Speech Audion, Processing, pp-129, 1993.
- [7] [7] Welling L. and Ney H., 'Formant Estimation of Speech Recognition', IEEE trans. Speech and Audio processing, pp-134. 1998.
- [8] Rabinar L.R., and Schafer R.W, 'Digital Processing of Speech Signal', Prentice Hall, Englewood Cliff, NJ, 1978.



Uzzal Sharma received MCA from IGNOU and PhD from Gauhati University. He is currently working as an Assistant Professor, Department of CSE &IT, DBCET, Assam Don Bosco University. He has got more than 10 years of teaching experience and around 3 years of research experience. He is the author of more than 20 research papers in different national and international journals and conferences. His research interest includes Speech Signal Processing and Computer Networks.