

ACOUSTIC REPRESENTATION OF GALO AND ADI VOWELS

Marpe Sora¹, P.H Talukdhar², Bomken Kamdak Bam³ and Firose A⁴

¹ Computer science & Engineering, Rajiv Gandhi University
Doimukh, Arunachal Pradesh, India

²Department of instrumentation & USIC, Gauhati university
Gauhati, Assam, India

³Computer science & Engineering, Rajiv Gandhi University
Doimukh, Arunachal Pradesh, India

⁴Computer science & Engineering, Rajiv Gandhi University
Doimukh, Arunachal Pradesh, India

Abstract

The vowel is the most distinguishing element of speech, feature vector extracted from the vowel have been considered for identifying the Language. In the present study, speech data is collected from the native speakers of Galo and Adii languages who are fluent in speaking and writing the language. Male and female speaker of age between 15 to 30 years, possessing a pleasant and a good voice quality are chosen to record the data. The recording is done one-by-one manner. The speakers were instructed to read each vowel naturally, without emotions and expression. They were asked to speak clearly and to keep their normal speaking rate and volume. To keep the recording consistent, both in phonetic and prosodic (within the framework of symbolic Prosody) terms, an expert in acoustic phonetics supervised the recording. The average duration of recording session was about 2 hours (2 recording session) for each speaker (Male & Female). We have recorded and used the native speakers data sets for analysis of the cepstral coefficients of vowel.

Key words: Prosody, Galo, Adii, cepstral.

I. INTRODUCTION

The Galo and Adii are one of the major tribes of Arunachal Pradesh India, lives in both the hilly and plain regions of Arunachal Pradesh. The languages in Arunachal Pradesh can broadly be classified into two groups: namely Abotani group and Non-Abotani (Buddhism). The majority of the tribes of the state belong to Abo Tani group. Galo and Adii belong to Tani group [1]. Their languages have many common words, accent etc. These language groups are very close both in syntax and semantics. The both tonal languages which has strong effects of intra-Tani contacts on the development of Tani languages. Both the languages they follow roman script and it has seven vowel each [2]. The Galo and Adii languages has seven vowels in which seventeen consonants [3] Galo and twenty three consonants of Adii. There is slight change in their representation. However both the tribe use same roman script for the writing purpose.

II. DETERMINATION OF MEL-FREQUENCY CEPSTRAL CO-EFFICIENT (MFCC)

The mel-cepstrum is a widely used parameter for speech recognition [5]. There are several methods that have been used to obtain the Mel-Frequency Cepstral Coefficient (MFCC). In the speech processing technology, The Mel-Frequency Cepstrum (MFC), is referred as the representation of short term power spectrum of a sound signal, based on a linear cosine transform of a log power spectrum on a non linear mel-scale of frequency. MFCC's are based on the known variation of the human ear's critical bandwidths with frequency. The speech signal is expressed in the Mel frequency scale for determining the phonetically important characteristics of speech. The following formula is used to compute the mels for a given frequency f , in Hz [4]:

$$mel(f) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \dots \dots (I)$$

As the mel spectrum coefficients are real numbers, they may be converted to the time domain using the **Discrete Cosine Transform (DTC)**. The MFCCs is calculated using the equation (II) [3]:

$$C_n = \frac{2}{N'} \sum_{k=1}^{N_f} X_k \cos \left(K_i \frac{2\pi}{N'} n \right); 1 \leq n \leq p \dots \dots (II)$$

Here K represents the number of mel cepstrum coefficient, which has been taken as 12 in the present study. C_0 , is excluded from the DCT as it represents the mean value of the input signal which carries little speaker specific information. For each speech frame of about 20ms with overlap, a set of mel-frequency cepstrum coefficients is computed. This set of coefficients is called an acoustic vector. To represent and recognize the speech characteristic of the speaker.

III. EXPERIMENTAL SETUP

The headset microphone was used for recording purpose and at the rate of 16 kHz sampling rate and file format is .wav. Distance between speaker and microphone is 8 inch.

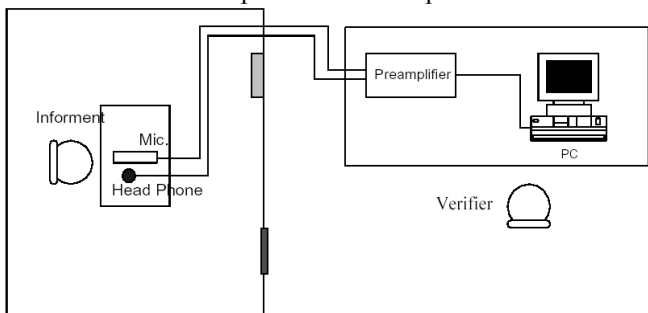


Figure 1.1 Recording setup

TABLE I: VARIATION IN ADII MALE VOWELS

Galo Vowels	Cmax	Cmin	Average	Range
\a\	0.8299	-0.8929	-0.0369	1.7228
\i\	0.8405	-0.7124	-0.0368	1.5529
\u\	0.6439	-0.7727	-0.0367	1.4167
\e\	0.8003	-0.9745	-0.0369	1.7748
\o\	0.9709	-1	-0.0365	1.9709
\v\	0.7310	-1	-0.0370	1.7310
\w\	0.4770	-0.5708	-0.0372	1.0478

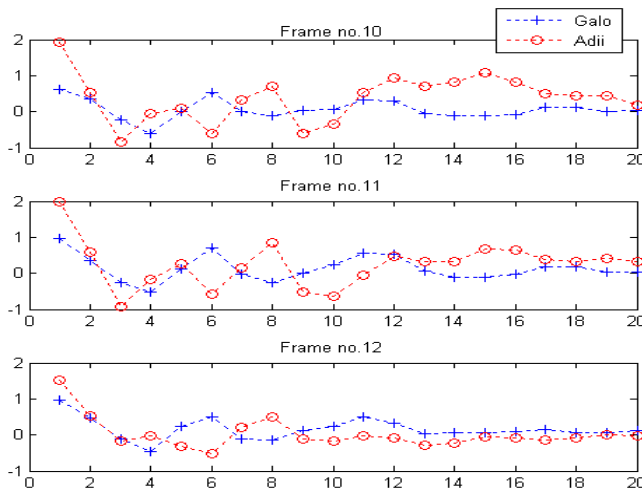
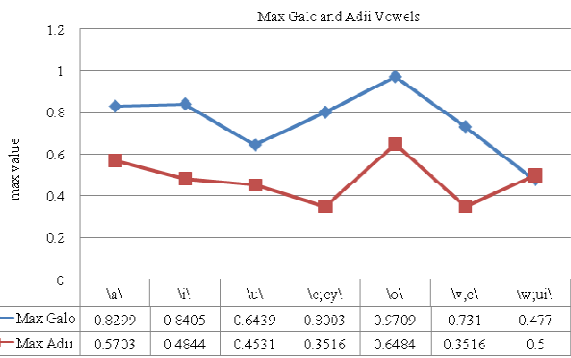


Fig 1. Vowel \a\ Galo and Adii

TABLE II: VARIATION IN GALO MALE VOWELS

Adii Vowels	Cmax	Cmin	Average	Range
\a\	0.5703	-0.4063	-1.7621e-004	0.9766
\i\	0.4844	-0.6094	-3.9193e-004	1.0938
\u\	0.4531	-0.5469	-1.3678e-004	1
\ey\	0.3516	-0.4375	-6.3073e-004	0.7891
\o\	0.6484	-0.7188	-1.7340e-004	1.3672
\e\	0.3516	-0.4375	-6.3073e-004	0.7891
\ui\	0.5000	-0.5625	-2.6926e-004	1.0625



In case of vowel \i\ the variation can be seen in the frame no.13,14 & 15.

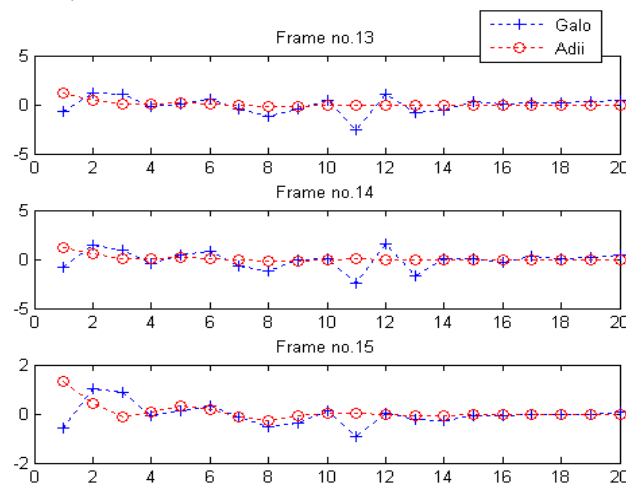
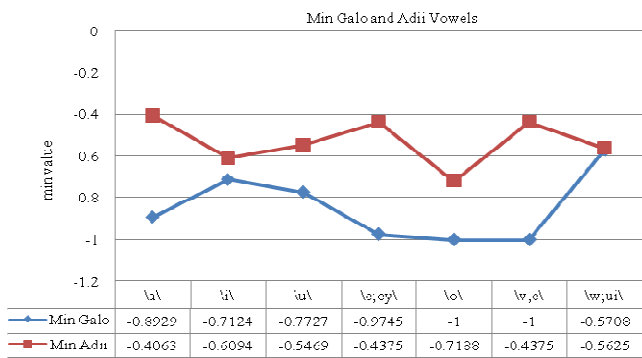


Fig1.1: Vowel \i\ Galo and Adii

In case of vowel Galo \u\ in both representation the variation can be seen in the frame no.13,14 & 15.

In the Vowel \a\ out of fifteen frames as per observation the frame no 10, 11 and 12 shows more dissimilarities in both Galo and Adii language in case of male speakers in both tribe, And other twelve frames shows almost similar values. Moreover in case of Adii more fluctuations are observed in MFCC.



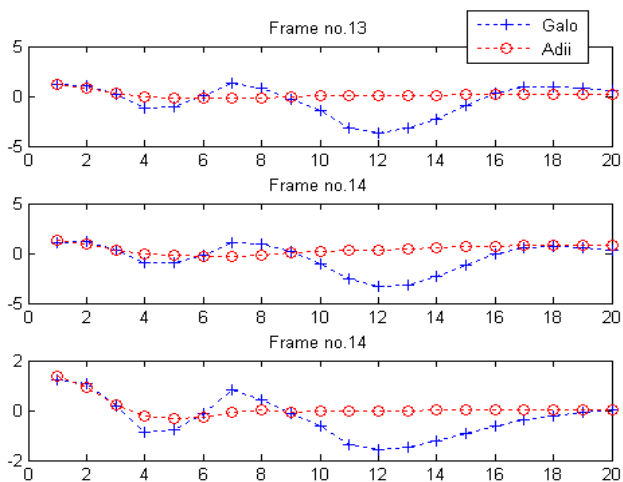


Fig1.2:Vowel 'u' Galo and Adii

In case of vowel Galo 'e' and equivalent representation in adii as 'ey' the variation can be seen in the frame no.10,11 & 12.

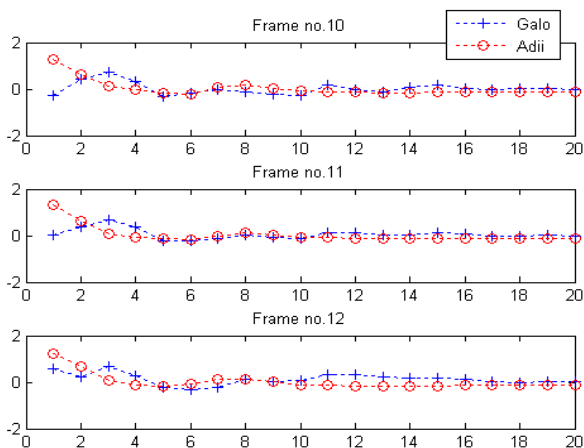


Fig1.3:Vowel 'e,ey' Galo and Adii

In case of vowel Galo 'o' the variation can be seen in the frame no.7,8 & 9.

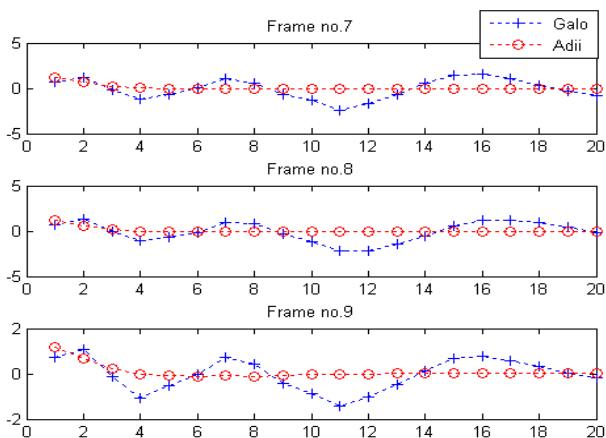


Fig1.4:Vowel 'o' Galo and Adii

In case of vowel Galo 'v' and equivalent representation in adii as 'e' the variation can be seen in the frame no.13,14 & 15

In case of vowel Galo 'w' and equivalent representation in adii as 'ui' the variation can be seen in the frame no.10,11 & 12.

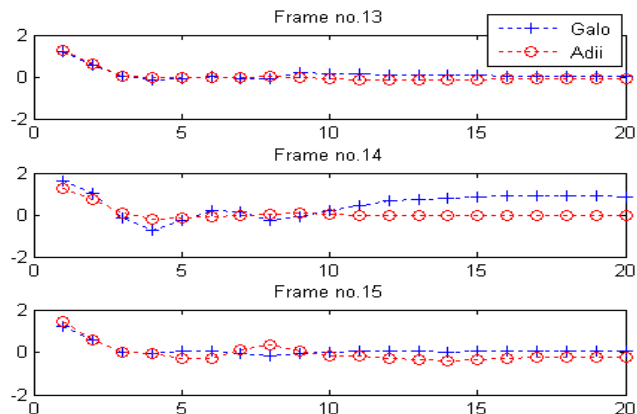


Fig1.5:Vowel 'v,e' Galo and Adii

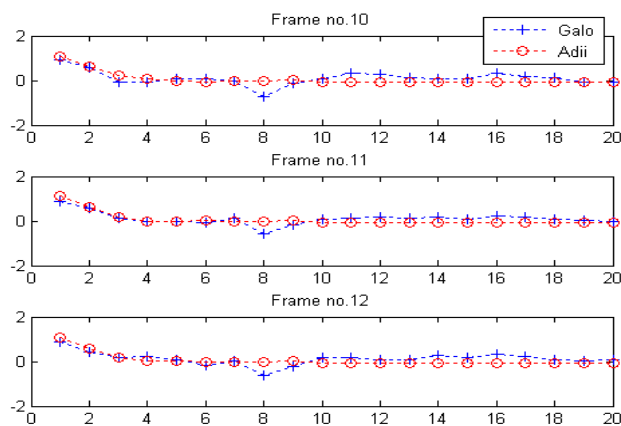


Fig1.6:Vowel 'w,ui' Galo and Adii

IV. VARIATIONS OF GALO AND ADI F2F VOWELS

The following given table shows the variations of max, min, average and range of the Galo and Adii Tribe languages.

TABLE III: VARIATION IN GALO FEMALE VOWELS

Adii Vowels	Cmax	Cmin	Average	Range
\a\	0.7344	-0.7891	-0.0015	1.5234
\i\	0.4375	-0.6328	-0.0013	1.0703
\u\	0.5703	-0.7109	-0.0051	1.2813
\ey\	0.4688	-0.4219	6.3247e-004	0.8906
\o\	0.7031	-0.5859	0.0013	1.2891
\e\	0.6953	-0.5234	2.3002e-005	1.2188
\ui\	0.4141	-0.5234	-1.2636e-004	0.9375

TABLE IV: VARIATION IN ADII FEMALE VOWELS

Galo Vowels	Cmax	Cmin	Average	Range
\a\	0.7462	-0.8013	-0.0362	1.5476
\i\	0.3627	-0.2515	-0.0358	0.6142
\u\	0.8307	-0.6755	-0.0360	1.5062
\e\	0.8385	-0.9621	-0.0361	1.8006
\o\	1.0000	-1	-0.0352	2.0000
\v\	1.0000	-1	-0.0382	2.0000
\w\	0.7003	-0.6946	-0.0362	1.3949

CONCLUSION

As per examine the properties of frame generated through matlab programme. The range of variations are more in case of male to male comparisons compare in case female to female comparisons. So as per analysis of 25 speaker’s male and 25 of female the larger variation are seen in case of male as per above tabulated parameter. This observation may be helpful in sex determination for both native speaker of Galo and Adii language.

REFERENCES

- [1] A Grammar of Nyishi language by “P.T Abraham”
- [2] Marpe Sora, Jyotismita Talukdar, and P. H. Talukdar “Formant Frequency and Cepstral Method Estimation of Galo Phonemes Using Acoustical Cues” *International Journal of Information and Electronics Engineering, Vol. 3, No. 1, January 2013*
- [3] Galo Agom Published by “Galo welfare society” May 2007
- [4] L. R. Rabiner and R. W. Schafer, *Theory and Application of Digital Speech Processing*, pp. 425-452, 2009.
- [5] Rabiner, L.R. and Schafer, R. W.; *"Digital Processing of Speech Signals"*, Englewood Clifts, N. J. Prentice Hill, 1978



Bomken Kamdak Bam obtained B.Tech form GCT Coimbatore under Bharathiar University Tamil Nadu and M.Tech from Rajiv Gandhi University. Presently he is Assistant professor in Rajiv Gandhi university, Department of Computer Science and Engineering Arunachal Pradesh. Her main research interests include signal and speech processing, software engineering and Computer Architecture.



Firose A received the Master of Engineering (ME). degree in Computer Science and Engineering from Anna University, Chennai in 2007 . Presently He is working as an Assistant Professor in Department of Computer Science and Engineering, Rajiv Gandhi University. His research interests include Networking, Network Security, Ad Hoc Networks, and Client/Server Technologies.



Marpe Sora obtained B.Tech form NERIST and M.Tech from Tezpur University. Presently he is Assistant professor in Rajiv Gandhi university, Department of Computer Science and Engineering Arunachal Pradesh. His main research interests include signal and speech processing, Data mining and HCI. He is member of RESONICS society under Gauhati University, Member of ISROSET(M-00000) International Scientific

Research organization for Science and Engineering Technology and member of IACSIT member(80346082) International association of Computer Science and information Technology



P.H. Talukdar Professor, Dept. of Instrumentation & USIC, GU (2006 to till date) Reader, Dept. of Instrumentation & USIC, GU (1996 to 2006) Assistant Professor, Dept. of MCA, Assam Engg. College (1991 to 1996) Assistant Professor, Dept. of Electronics & Telecommunication Engineering., Assam Engineering College (1986 to 1991) Lecturer, Dept. Of Physics, Cotton College, Guwahati (1982 to 1986). Lecturer, Dept. of Physics, Bajali College (

1981 to 1982) Lecturer, Dept. of Physics, D K College, Mirza (1979 to 1981) Scientific Assistant, Defence R & D Organization, Dehradun (1976 to 1979).He is author of six technical books. His main research interests include signal and speech processing.