International Journal of Scientific Research in Computer Science and Engineering

Research Paper
Volume-1, Issue-1, Jan- Feb-2013
Available online at www.isroset.org

ISSN: 2320 - 7639

# Word level detection of Galo and Adi language using acoustical cues

Marpe Sora[#1], Jyotismita Talukdhar[#2], Swanirbhar Majumder[#3], P.H Talukdhar[#4],U.sharmah[#5]

[#1]Department of computer Science and Engineering/ Rajiv Gandhi university /India, msorait@yahoo.co.in
[#2]Department of Computer Science/ Asian Institute of Technology /Thailand, jyotismita4@gmail.com
[#3]Department of Electronics/ North eastern Institute of Technology /India, swanirbhar@gmail.com
[#4] Departement of Instrumentation & USIC /Guwahati University/India, phtassam@gmail.com
[#5]Departement of Computer Science NERIM/Dibrugarh university/India, uzzal_sh@rediffmail.com

***Abstract*— Word level Speech recognition is the process of automatic extracting and determining linguistic information Conveyed by a speech wave using computers.A successful speech detection system can helpful in many applications. Study of acoustical cues of Galo and Adi language is main theme. The Galo and Adi tribe belongs to the Sino-Tibetan family of languages. The state of Arunachal Pradesh has evolved a conspicuous dialect that distinguishes them from the rest of the tribes. A further classification can be based on the "dialect" of the language. The variations of Galo and Adi spoken in these tribe differ from each other in intonation, sentence formation and word usage.**

***Keywords*— EMD, Abotani, Galo, Adii**

## I. INTRODUCTION

All standard paper The Galo and Adi language is a Tibeto-Burman language of the Tani group. They were formerly known as Abhors. The Galo and Adi speaking constitute 21% total population of Arunachal Pradesh . In the Arunachal Pradesh there are total 25 major tribes and almost 110 sub tribes. They have high degree of mutual intelligibility is found across the languages spoken by the Adis, the Apatanis, the Galos, the Hill Miris, the Nyishis and the Tagins. Moreover they share many characteristic features in their cultural code and trace their ancestry from a common forefather, namely Abotani. Hence, the language spoken by them can rightly be given generic name –Tani language [1].Besides the languages in Arunachal Pradesh can broadly classified into two groups: namely Abotani group and Non-Abotani(Buddhism). The majority of the tribes of the state belong to Abo Tani group and Galo and Adi belong to Tani group this two tribe are very close both in syntax and semantics. Speakers of Galo inhabitant the districts of West Siang, East Siang and south eastern side of Upper Subansiri, as well as in some small pockets in Lohit, Lower Dibang Valley and Changlang Districts of Arunachal Pradesh in the North-East Indian Himalaya. The Adis spread over the East, west and upper siang districts of Arunachal Pradesh. Both tribes have contributed significantly in enriching the ethnic tradition in the state. The studies are more prominent today due to large changes in society..

## II. EMD BASED APPROACH

First, The main idea of EMD method is to decompose time series signal into some intrinsic mode functions (IMF) and one residue. The name 'intrinsic mode function' is adopted because it represents the oscillation mode imbedded in the data. An intrinsic mode function (IMF) is a function that satisfies two conditions: (1) in the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

*The Algorithm EMD process is as follow [2]:*

1) Detect the extrema (Both minima and maxima) of original signal $s(t)$.
2) Generate the upper and lower envelops $h(t)$ and $l(t)$ respectively by connecting the maxima and minima separately with cubic spline interpolation.
3) Determine the local mean as $m_1(t) = [\, h(t) + l(t) \,]/2$
4) The upper and lower envelopes should encompass all the data to designate the mean $m_1(t)$ of two envlopes. The difference between the data $s(t)$ and $m_1(t)$ is first component $h_1(t)$,i.e

$$h_1(t) = s(t) - m_1(t) \quad (1)$$

Ideally, $h_1(t)$ should satisfy the two definitions and is an IMF. However, in practice, all the conditions of an IMF cannot be satisfied. So above the shifting process is repeated. $h_1(t)$ is viewed as new original signal. The shifting process is repeated until first IMF component $c_1(t)$ satisfies the above two definitions.

5) Then the difference between the data $s(t)$ and $c_1(t)$ is the residue $r_1(t)$:

$$s(t) - c_1(t) = r_1(t) \quad (2)$$

$r_1(t)$ is then treated as new data and subjected to the same shifting process as above to obtain. Thus procedure are repeated to obtain all the subsequent

$c_2(t), r_2(t), c_3(t), r_3(t)\ c_3(t)$ ……., $c_n(t)\ r_n(t)$. The shifting process will stop if $r_n(t)$ becomes so small that it is

less than a predetermined value of consequence and a monotonic function, from which no more IMF can be extracted.

$$r_1(t) - c_2(t) = r_3(t); \ldots . r_{n-1}(t) - c_n(t) = r_n(t) \quad (3)$$

So the original signal can be computed as the sum of the IMF components plus the final residue:

$$s(t) = \sum_{n=1}^{n} c_n(t) + r_n(t) \quad (4)$$

Where $n$ is the num

ber of IMFs and $r_n(t)$ is final residue.

It is a new nonlinear technique has recently pioneered by N.E Huang *et al*. for adaptively representing non stationary signals as sums of zero-mean AMFM components [3] However interestingly the useful decomposition outcomes, has found a vast number of diverse applications such as biomedical[4].

TABLE I. 1GALO MALE AND FEMALE WORDS AVG. RANGE OF EMD VARIATION

| Galo Word Uttered | English Word | Avg.range of variation Male | Avg.range of variation Female |
|---|---|---|---|
| AO | Son | 0.36055 | 0.2843 |
| ALAK | Hand | 0.18625 | 0.3241 |
| ALE | Leg | 0.31975 | 0.27685 |
| DONYI | Sun | 0.1277 | 0.21515 |
| DENAM | Flying | 0.1876 | 0.50275 |
| DUNAM | Sitting | 0.12565 | 0.2767 |
| EMME | Firing | 0.25035 | 0.27735 |
| GOKNAM | Calling | 0.1685 | 0.20445 |
| HO-II | Cow | 0.2047 | 0.2001 |
| JEMNAM | Chewing | 0.2796 | 0.38715 |
| KAJU | Let go | 0.29915 | 0.3899 |
| KATO | Seen | 0.34925 | 0.43025 |
| LAGNE | Thumb | 0.2207 | 0.48905 |
| LOBO | Summer | 0.1549 | 0.148 |
| MENAM | Lie | 0.1018 | 0.20095 |
| MOMIIN | Fight | 0.0794 | 0.1694 |
| NAMNAM | Smell | 0.0769 | 0.19365 |
| NO | You | 0.07295 | 0.19975 |
| NENAM | Envy | 0.1185 | 0.3566 |
| NYIIRNAM | To laugh | 0.06625 | 0.3971 |
| NGO | I am | 0.07415 | 0.1892 |
| POLO | Moon | 0.1394 | 0.53935 |
| RAMNAM | Fever | 0.11225 | 0.48455 |
| RENAM | Buy | 0.12505 | 0.10825 |
| ROKOM | Morning | 0.087 | 0.35755 |
| TADOK | Necklace | 0.30845 | 0.73445 |
| TARUK | Ant | 0.28815 | 0.6828 |
| TIINAM | Drink | 0.18905 | 0.5828 |
| TARIIK | Date | 0.2506 | 0.52225 |
| YANA | Rotten | 0.28955 | 0.49225 |
| YENAM | Live | 0.117 | 0.3456 |
| YUPNA | Sleeping | 0.20465 | 0.2889 |
| DAKANAM | Stand | 0.29795 | 0.42655 |
| DIIKNAM | Burning | 0.06235 | 0.51635 |
| DONAM | Food | 0.12805 | 0.25785 |
| GENAM | To carry | 0.2511 | 0.20265 |
| JIINA | Healthy | 0.2882 | 0.1955 |

ZTABLE II. ADII MALE AND FEMALE WORDS AVG. RANGE OF EMD VARIATION

| Adi Word Uttered | English Word | Avg.range of variation Male | Avg.range of variation Female |
|---|---|---|---|
| AO | Son | 0.1924 | 0.3963 |
| ALAK | Hand | 0.0738 | 0.3804 |
| ALE | Leg | 0.1223 | 0.46465 |
| DONYI | Sun | 0.08235 | 0.2136 |
| DENAM | Flying | 0.0782 | 0.597 |
| DUNAM | Sitting | 0.04065 | 0.1878 |
| EMME | Firing | 0.0484 | 0.3117 |
| GOKNAM | Calling | 0.09685 | 0.16175 |
| HO-II | Cow | 0.05605 | 0.2954 |
| JEMNAM | Chewing | 0.05755 | 0.26225 |
| KAJU | Let go | 0.07635 | 0.47655 |
| KATO | Seen | 0.08375 | 0.27855 |
| LAGNE | Thumb | 0.04215 | 0.188 |
| LOBO | Summer | 0.02095 | 0.25415 |
| MENAM | Lie | 0.03375 | 0.3754 |
| MOMIIN | Fight | 0.0334 | 0.29555 |
| NAMNAM | Smell | 0.0479 | 0.2268 |
| NO | You | 0.02955 | 0.30485 |
| NENAM | Envy | 0.0358 | 0.2653 |
| NYIIRNAM | To laugh | 0.03725 | 0.2784 |
| NGO | I am | 0.03125 | 0.1969 |
| POLO | Moon | 0.11675 | 0.32785 |
| RAMNAM | Fever | 0.1126 | 0.16065 |
| RENAM | Buy | 0.057 | 0.23255 |
| ROKOM | Morning | 0.057 | 0.23535 |
| TADOK | Necklace | 0.11995 | 0.35885 |

| | | | |
|---|---|---|---|
| TARUK | Ant | 0.17755 | 0.44905 |
| TIINAM | Drink | 0.18195 | 0.36095 |
| TARIIK | Date | 0.29035 | 0.33445 |
| YANA | Rotten | 0.2036 | 0.2454 |
| YENAM | Live | 0.1567 | 0.28755 |
| YUPNA | Sleeping | 0.139 | 0.43535 |
| DAKANAM | Stand | 0.12295 | 0.3557 |
| DIIKNAM | Burning | 0.168 | 0.1967 |
| DONAM | Food | 0.1255 | 0.2925 |
| GENAM | To carry | 0.08625 | 0.1514 |
| JIINA | Healthy | 0.15165 | 0.1819 |

The IMF1, IMF2, IMF3 and time graph are shown below of the word Alak. It is observed that higher order IMFs contain lower frequency oscillations than that of lower order IMF [5].

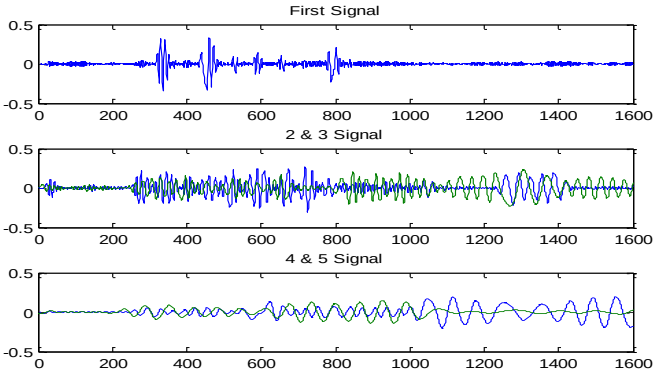


Figure 1. Word Alak

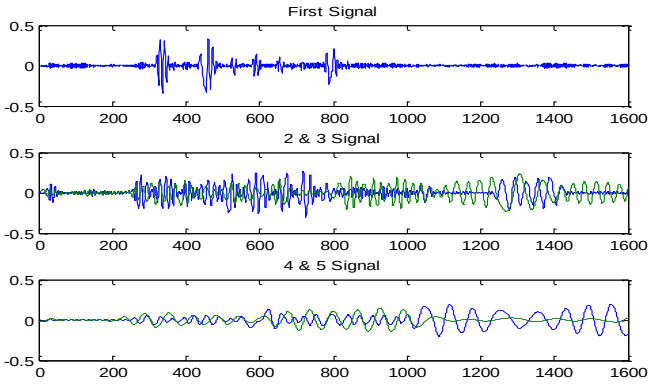


Figure 2. Word Ale

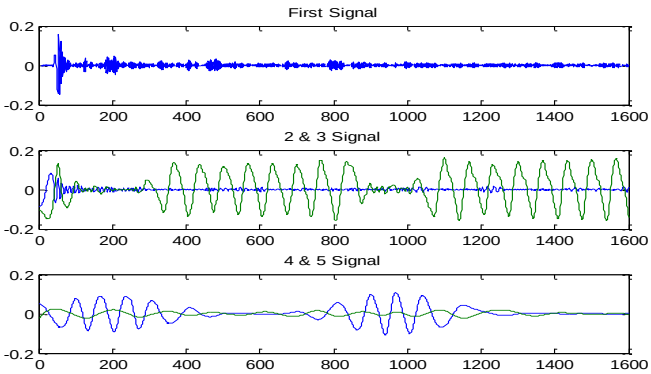


Figure 3. Word Dunam

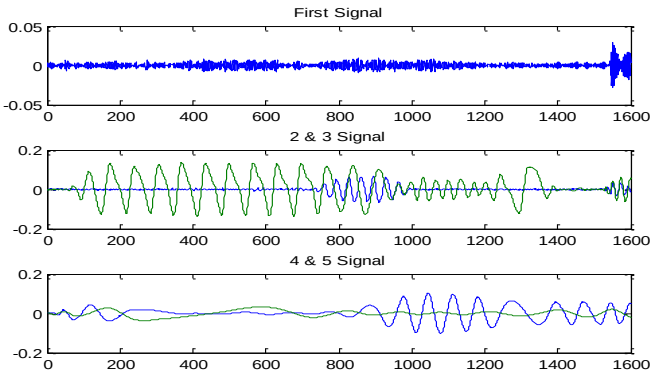


Figure 4. Word Goknam

## III. METHODOLOGY

The data used for EMD approach is primary data recoding was taken in the closed room of 37 similar words and having same meaning in which equivalent English meaning are shown in the table. As it is shown in the table I & II, by which data are obtained for analysis. We have recorded 20 male and 20 female and recoding time is 10hrs 15 minutes. For recording purpose we have taken age between 15 to 30 years. Who are fluent, command and wring skill in their language.

## IV. OBSERVATIONS AND CONCLUSIONS

As per the table I & II. The EMD based analysis of Galo and Adi words the following observations were made-Significant variation are observed among the Galo words as shown in Table I. The variation is found to be maximum with respect to word /Ao/ and minimum corresponding to word /alak/, in case of male speakers. Similarly, for female Galo speakers, the maximum variation of autocorrelation is found corresponding to word /Ao/ and minimum in case of /Alak/.

In case of Adi words, i.e., /Ao/, /Alak/ for both male and female speakers the range of variation of the (Table II) is found to be maximum in case of male speakers with respect to word /Ao/ and minimum with respect to vowel /Alak/. In case of female speaker, the maximum variation of autocorrelation is found in case of word /Ao/ and minimum with respect to vowel /Alak/.

The variation of the EMD for male is very irregular in contrast to the stable variation of female EMD. The same phenomenon is also observed in case of Adi word also, but in this case the frame numbers are different. This observation may be helpful in sex determination for both Galo and Adi speakers.

The range of variation of EMD for Galo and Adi male is found within the range of 0.36055>EMDGalo>0.06235 and 0.29035>EMDAdi>0.02095 respectively. The range of variation for female is found 0.73445>EMDGalo>0.10825 and 0.597> EMDAdi>0.1514 i.e. the variation of EMD analysis for Galo words is more (Male-0.2982; Female-0.6262)with respect to the Adi words (Male-0.2694;Female-0.4456) i.e., the former is stable as compared to the latter. The results reported here are believed to provide with new insights on EMD and its use.

## REFERENCES

[1]. Galo Agom published by "Galo Welfare Society" Itanagar , May **2007**.

[2]. Y.Liu,H.Yang and H.Zhou "Speaker identification based on EMD" IC-NIDC Proceeding pp.**808-812,2009**.

[3]. N.E. Huang, Z. Shen, S.R. Long, M.L. Wu,H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, "The empirical mode Decomposition and Hilbert spectrum for nonlinear and non stationary time series analysis," *Proc. Roy. Soc.London A*,Vol. 454, pp. **903–995, 1998**.

[4]. Y. Zhang, Y. Gao, L. Wang, J. Chen, and X. Shi, "The removal of wall components in Doppler ultrasound signals by using the empirical mode decomposition algorithm," IEEE Trans. Biomed. Eng., vol. 9, pp.**1631–1642**, Sep. **2007**.

[5]. Md. Khademul Islam Molla, Keikichi Hirose, Nobuaki Minematsu, and Md. Kamrul Hasan voiced/unvoiced detection of speech signals using empirical mode decomposition model" International Conference on Information and Communication Technology ICICT 2007, 7-9 March **2007.**

## AUTHIR PROFILE

**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. Presently he is also pursuing his PhD from Guwahati University.

**Joytismita Talukdhar** obtained B.Tech form Guwahati and perusing Master in Computer Science from Asian Institute of Technology Bangkok Thailand. Her main research interests include signal and speech processing.

**Swanirbhar Majumder** did his ITI, Diploma & B.Tech from Dept of ECE, NERIST, Arunachal Pradesh under NEHU, Shillong in the years 2000, 2002 ans 2004 receptively followed by M.tech from the Dept of Applied Physics, University of Calcutta. in 2006. He Joined NERIST as Asst Professor in July 2006 and has been serving there till date. His interests include Image Processing, Speech and Biomedical Signal Processing, Embedded Systems and ANN. He has published more than 30 international papers in journals, conferences and book chapters. Presently he is also pursuing his PhD from Dept of ETCE, Jadvpur University.

**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 to1979 ).He is author of six technical books. His main research interests include signal and speech processing.

**Dr. Uzzal Sharma** obtained MCA form IGNOU and Ph.D from Gauhati University. Presently he is Assistant professor in the Department of Computer Science NERIM, Guwahati. His research areas include Speech and Signal processing and Computer Networks.