Study of Phoneme Level Spectral-Temporal Difference between Tribal (Paniya) and Malayalam Language

Ponnu Mathew Thomas[1], Riyas K.S.[2], Jayan A.R.[3]
[1, 2, 3] Department of ECE
[1] Govt. Engineering College Wayanad, Thalappuzha-670644
[2] Rajiv Gandhi Institute of Technology, Govt. Engineering College, Velloor- 686501, Pampady
[3] Govt. Engineering College Sreekrishnapuram, Mannampatta- 679513, Palakkad, Kerala, India

Abstract- *The ultimate goal of speech processing research is to come up with a system by which a machine can recognize, understand and produce speech just like a human being. For the development of any such system requires a detailed study of the various spectral and temporal features of speech sounds in the corresponding language. This paper is a comparative study between the two Dravidian languages - Tribal (Paniya) and Malayalam using time domain and frequency domain features. Here, speech analysis is being done on a phoneme level basis on isolated word database. The various phones are clubbed into broad phoneme classes like vowels, nasals, stops, approximants, fricatives and silence according to the characteristics shown by them. A detailed study is carried out to identify features useful for broad phoneme classification by analyzing sounds from different classes. The features identified useful for discriminating Malayalam and Tribal (Paniya) phonemes are phoneme duration, energy, spectral flatness, spectral centroid, pitch and spectral moments (mean, variance, kurtosis and skewness) of fundamental frequency.*

Keywords- Phoneme, Duration, Energy, Spectral Centroid, Spectral Moments.

I. INTRODUCTION

Speech is the efficient way through which humans communicate with each other. The efficiency of speech recognition systems have readily improved due to the advancement in technology, development of better acoustic models, new feature extraction algorithms, better NLP techniques, DSP tools, as well as hardware and software tools and a steady progress is observed in the recognition of complex speech patterns. On comparing the speed with which humans can recognize speech the performance of the systems is still low, especially in cases of adverse situations like background noise, disturbances and other degradable conditions that affect the quality of speech signals. This created a considerable performance gap between humans and machines. So there is a scope for more research and development in this area so as to facilitate a computer to achieve the speech recognition ability identical to that of humans.

Speech with its own diversity and inherent variations is one of the principal and ultimate techniques in recognizing an individual through the process of verbal signs. A lot of researchhas been noticed in the languages like English, Bengali, Hindi, Tamil, Chinese, Arabic etc. But only a few have been reported in Tribal and Malayalam languages. This paper involves the analysis of two Dravidian languages - Malayalam and Tribal (Paniya). This study has a great importance and is one of the challenging areas of speech researches.

'Malayalam' is the official language of Kerala and one among the twenty two scheduled languages of India. This language is spoken by the people of Kerala and in the union territories of Lakshadweep and Puducherry. It is the youngest and toughest among the developed languages of the Dravidian family. 'Panyah' is a language of the tribal sect (Paniya) who resides mainly in Kerala. This is one among the Dravidian languages, also known as Pania and Paniyan. The people of this sect are living according to traditional cultures like any other tribal folks. Major section of Paniyans in Kerala reside in the districts of Wayanad, Kannur, Kozhikode and Malappuram and to the west of the Nilgiri Hills in Tamil Nadu. This language like other tribal languages also has no literary specifications.

II. OVERVIEW OF THE STUDY

Speech perception is the process by which the speech sounds of any language is heard, interpreted and understood. It is the mapping done perceptually from the speech signal which is greatly variable to a linguistic representation in the form of phonemes, diphones, syllables, words etc. Phoneme or phone is the smallest unit of sound that distinguish one word from another in a particular language. Most languages have about 20-40 phones that give an alphabet of sounds to specify uniquely the words in a language.

A. Analysis of Broad Phoneme Classes of Speech Sounds

The broad phoneme classes of Malayalam language, their characters and corresponding IPA symbols are given in the figure 1.

SL No.	Broad Phoneme Class	Symbols Used To Represent The Class	IPA Symbols Of Phonemes Within Each Class	Corresponding Malayalam Characters
1	Vowels	V	a i u e o a: i: u: e: o:	അ ഇ ഉ എ ഒ ആ ഈ ഊ ഏ ഓ
2	Nasal	N	m ɱ n ɳ ɲ ŋ	മ ന ണ ഞ ങ
3	Stop	P	b ɖ d̪ dʒ g bʰ ɖʰ d̪ʰ dʒʰ gʰ p ʈ t̪ tʃ k pʰ ʈʰ t̪ʰ tʃʰ kʰ t	ബ ഢ ഡ ജ ഗ ഭ ധ ഝ ഥ ഘ പ ത ട ച ക ഫ ഥ ഠ ഛ ഖ ററ
4	Fricative	F	f s ʂ ʃ h	ഫ സ ഷ ശ ഹ
5	Approximant	A	ʋ ɻ j l ɭ r ɾ	വ ഴ യ ല ള ര റ
6	Silence	S	-	-

Fig.1. Broad Phoneme Classes with Malayalam characters

The total sound units were classified in to six groups as shown in the figure. For grouping the input speech units in to any one of the six broad phoneme classes, some particular features were needed for each class. For this an analysis study was conducted to check whether there are some special features available for each class. The phonemes analyzed in the work are given in table 1.

(i) Voiced and unvoiced sounds can be easily recognized. Unvoiced sounds have high zero crossing rate, pitch will be zero and they are aperiodic in nature. Whereas voiced sounds have lower zero crossing rates, their pitch will not be equal to zero, and are almost periodic signals.

TABLE I : BROAD PHONEME CLASSES ANALYSED IN THE WORK

	Vowel	*Plosive*	*Nasal*	*Approximant*	*Silence*
Voiced	a, a:, o, o:, u, i	d,ch	m,n, mm	v, l, r	-
Unvoiced	-	p, k, t	-	-	-
Presence of Burst	-	Yes	-	-	-

(ii) Vowels are voiced sounds with periodic amplitude and high intensity. They have longer duration of about 60ms in the utterance.

(iii) Nasals are also periodic signal with lesser amplitude than vowels. They are also voiced sounds. If their formants are analyzed, we can see that the first formant frequency of nasal is weaker. By evaluating only the formant frequencies we can separate these classes: vowels and nasals.

(iv) Plosives or stops are an oral occlusive. It is a consonant sound made by preventing the flow of air in the vocal tract. Plosive sounds are of two types: voiced and unvoiced. The main peculiarity of plosives is the presence of burst in the signal.

(v) Approximants are the sounds where the articulators come close to each other but neither narrowly enough nor with enough articulatory precision to create a turbulent flow of air. Therefore, approximants lie between vowels and fricatives where the vowels produce no turbulence and fricatives produce turbulent airstream. The examination of the waveform of /v/ of the word /vaal/ shows periodic large peaks. Superimposed on these are small irregular aperiodic peaks. This is actually the superimposition of weak fricative over strong vowel sound.

B. Analysis of Spectral and Temporal Features of Speech Sounds

A study was conducted to identify the features for dif-ferentiating Malayalam and Paniya phonemes. Various temporal and spectral features explored include the following: (i)Voicing information (ii)Phoneme duration (iii)Short Time Energy (iv)Spectral Flatness (v)Spectral Centroid (vi)Pitch and (vii)Spectral Moments (including spectral mean, variance, skewness and kurtosis).

(i) Voicing Information: Voicing information helps to find out whether the speech frame is voiced or unvoiced. This information is obtained with the help of Wavesurfer's pitch contour.

(ii) Phoneme duration: Phonemes are the language category in which different phones are grouped as same. Phoneme duration is a time domain parameter. It is the length of a particular phoneme when uttered in speech.

(iii) Short Time Energy: Energy is the squaring operation done on the speech sample. High energy indicates high amplitude of speech waveform. It is a parameter which is simple to calculate. The high variation in amplitude between voiced and unvoiced speech and small variation between phonemes with different manners of articulation

permit speech segmentations based on energy.

(iv) Spectral flatness: Spectral flatness is a feature used to characterize the spectrum of an audio signal. It is also called Tonality coefficient. It provides a way to assess how tone-like a sound is, as opposed to being noise-like. The graph of the spectrum would appear relatively flat if the spectral flatness is high and close to that of white noise (value 1.0). The graph of the spectrum would be spiky if the spectral flatness is low and close to that of pure tone (value 0.0).

(v) Spectral Centric: Spectral Centric is a frequency domain feature. It is a measure used in DSP to characterize a sound spectrum. It shows where the center of mass of the audio spectrum lies. Spectral centric models sharpness of a sound. Higher the value of spectral centric, brighter will be the sound.

(vi) Pitch: Pitch is the fundamental frequency of vibration of vocal folds. They vibrate quasi- periodically only for voiced sounds. Generally it is higher for females in the range of 165-255 Hz than males which are in the range of 85-180 Hz. Pitch or F0 is the main acoustic cue to stress and intonation in speech. It is important in the phoneme identification of tone languages.

Spectral Moments: Spectral moment analysis (SMA) is generally used to analyze the patterns of acoustic energy found in a number of different speech sounds. The aim is to quantitatively describe the patterns of spectral energy within the band of noise characteristic of obstruent sounds. SMA uses numerical values to describe the spectral energy of a speech sound within a static window or period of time. Four primary spectral moments are the spectral mean, spectral variance, spectral skewness, and spectral kurtosis. The spectral mean is the energy distribution averaged over a specific part of the spectrum. The second spectral moment is the spectral variance. It denotes the deviation of frequencies represented in the spectrum, or how much these frequencies vary with respect to the mean. The third moment, spectral skewness describes the asymmetry or skewness of the noise energy distribution; it is also called spectral tilt. The fourth moment, spectral kurtosis suggests the peakedness of the spectral distribution.

III. METHODOLOGY

The various steps done for the comparative study of the two languages on the phoneme level basis are as follows.

A. Database Collection

Speech corpus in the form of isolated words was collected from the speakers of Malayalam and Tribal (Paniya) language. It was recorded under lab environment with the help of microphone and digital audio editing software Sony Sound Forge in Mattoli studio at Dwaraka, Wayanad. A set of 10 similar words each of the two languages were selected for the study. The words (Malayalam/Paniyah) are as follows: paal/paalu, vaal/vaalu, thol/tholu, nool/noolu, pakal/pakalu, manal/manalu, kadal/kadalu, thummal/thummalu, chithal/chithalu and ural/uralu. Each of these words was repeated 5 times. The dataset was collected from equal number of males and females of both languages. There was a total of 12 speakers (3 in each group). Thus, the study consisted of 600 utterances.

B. Pre-processing

Pre-processing of speech signals is the important part in the development of an efficient and robust speech recognition system after dataset creation. Various types of degradation components like environmental noise, background noise, effects from recording hardware, reverberation and disturbances interferes with the speech samples. This affects the quality of the audio signals. Furthermore, due to the variations in the speech signal, the sound signals that are similar visually may not produce similar sounds perceptually. Hence, pre-processing of the signal before the feature extraction has a great relevance on the performance of the speech recognition system. In this work, pre-processing involves the removal of noise and silence region preceding ad succeeding a word. This is done with the help of the audio editor tool called Audacity. The sound spectrum was converted from the stereo format to mono using the same tool.

C. Transcription

Transription of the words into phonemes is the main step involved in the work. The manual labeling of the collected read speech according the broad phoneme classes was themost tedious task. The issue related to the manual transcription mainly includes time consumption. For the manual phoneme level transcription of one minute speech data it takes about 30 hours. Since we are doing manual transcription it was prone to have mistaken in it. Moreover the transcriptions are done based on the sound heard. So it varies between people to person. For example one may label the sound to be a voiced stop and another person may label the same sound to an unvoiced stop. In this work, transcription is done manually using audio editor tool Wavesurfer under the option transcription.

D. Feature Extraction and Analysis

Feature extraction has a major role in parting one type of speech from the other type. It is an important part of ASR systems. Feature extraction is a technique of data reduction as it changes the input speech signal into a small set of parameters where the frequency domain (spectral) as well as time domain (temporal) characteristics of the speech signal are preserved. It also rejects the unwanted information from the signal. Feature extraction produces feature vectors that may be time domain or frequency domain characteristics depending on the technique used. Hence, feature extraction transforms the speech signal into a form suitable for classification. In the work, the various spectral and temporal features are extracted using the programming language Matlab.

E. Comparative Study

The final section of the thesis is the comparative study between the feature sets of the two languages- Malayalam and Tribal. The various temporal and spectral features were analyzed and the differences were noted. The remarkable differences shown in the feature set can be used in discriminating Malayalam and Tribal (Paniya) phonemes and can be used in classification systems. The pattern followed by the features of each language was also noted and studied in detail.

IV. RESULT AND DISCUSSIONS

The observation and results noted in each of the broad phoneme class for the selected spectro-temporal features are given below.

A. Phoneme Duration

Duration of the phonemes (both consonants and vowels) is a notable feature that distinguishes between Tribal (Paniya) and Malayalam phonemes. It is observed that the phoneme duration is a feature that depends on the position of the phone in the word.

Consonant Duration: The duration in uttering the plosive sound when it is present at the beginning of the word is more for Tribal than Malayalam speakers. This is seen in phones /th/ in the word /thol/ and /p/ in the word /paal/. When the plosive sound is found somewhere in the middle of the word like /k/ in the word /pakal/ and /d/ in the word /kadal/, no such differences were noted. This is shown in figure 2.

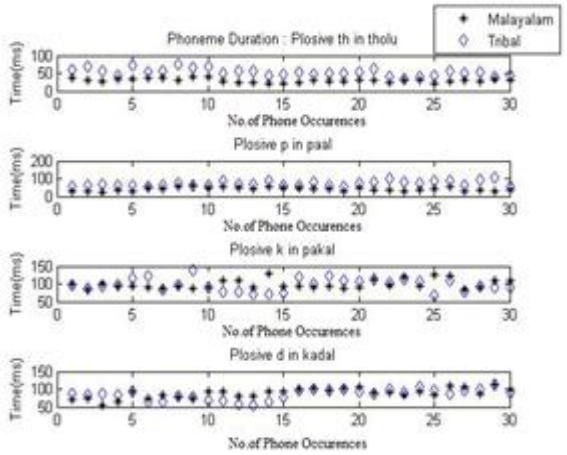

Fig.2. Duration of plosive sounds

The duration in uttering the nasal sound when it is present at the beginning of the word is more for Malayalam speakerthan Tribal speakers. This is seen in phones /n/ in the word /nool/ and /m/ in the word /manal/. When the nasal sound is found somewhere in the middle of the word like /n/ in the word /manal/ and /mm/ in the word /thummal/, no such differences were noted. This is shown in figure 3.

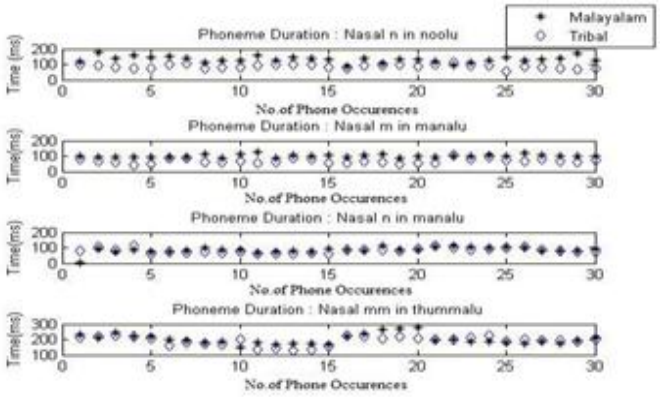

Fig.3. Duration of nasal sounds

The duration in uttering the approximants when it is present at the beginning or end of the word is more for Malayalam speaker than Tribal speakers. This is seen in phones /v/ in the word /vaal/ and /l/ in the word /paal/. When the approximant sound is found somewhere in the middle of the word like /r/ in the word /ural/, no such differences were noted. This is shown in figure 4.

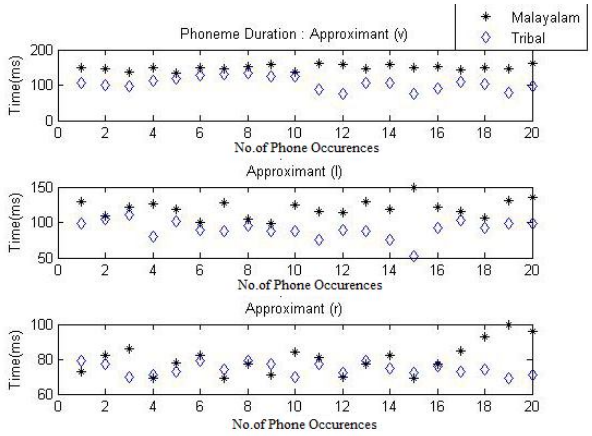

Fig.4. Duration of approximant sounds

Vowel Duration: The duration in uttering the vowels irrespective of the position of the phone in the word is more for Tribal than Malayalam speakers. This is seen in phones /a/ in the word /kadal/, /aa/ in the word /vaal/, /o/ in /thol/, /oo/ in /nool/, /u/ in /ural/ and /i/ in /chithal/. This is shown in figure 5.

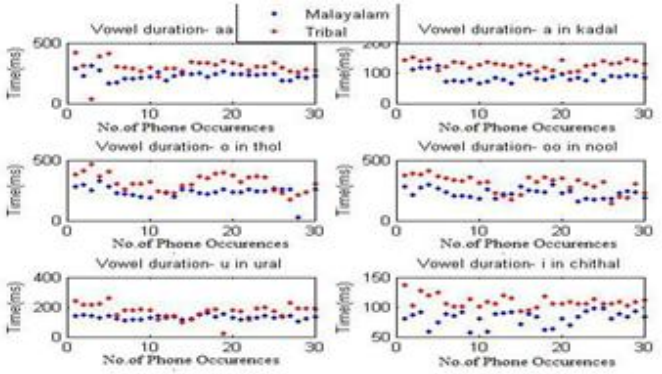

Fig. 5. Duration of vowels

Duration of word: Considering the duration of the complete word, it is found to be more for Tribal than Malayalam speaker. This is due to the fact that Paniyans show a tendency to prolong the vowel sounds.

B. Energy

Energy is a time domain parameter. Energy is reflected in the amplitude of the speech signal. This means that higher theamplitude of the speech waveform, higher will be its energy. Tribal speakers use high energy in uttering the stop (plosives) sounds compared to Malayalam speaker in uttering thesame phone which is shown in figure 6.

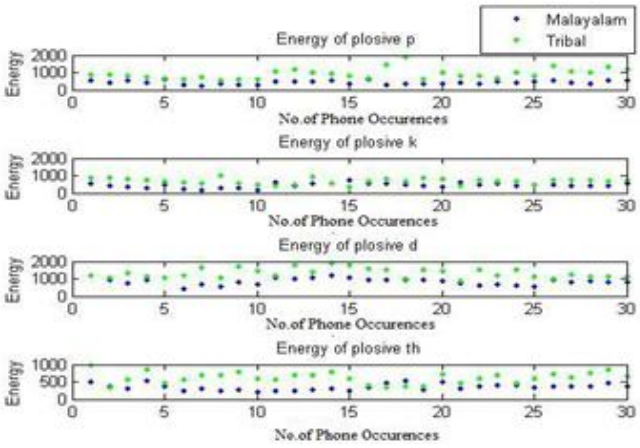

Fig.6. Energy of Malayalam and Tribal speakers in uttering plosive sounds

Energy of the Malayalam speaker is more when the nasal sounds are spoken. This is shown in figure 7.

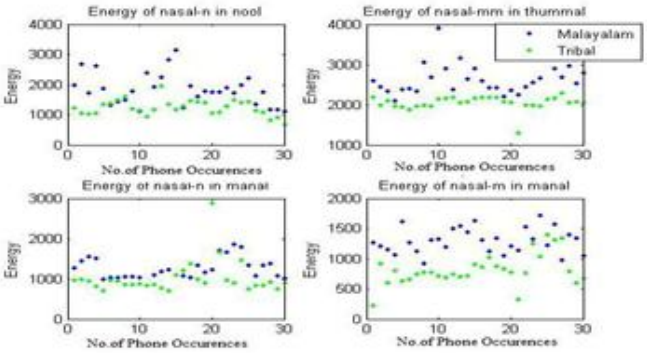

Fig.7. Energy of Malayalam and Tribal speakers in uttering nasal sounds

Energy of the Tribal speaker is found to be more in uttering vowels than Malayalam speakers as shown in figure 8.

Malayalam and Tribal speakers have energy in the same range in uttering approximants.

Energy of word: Considering the energy of speaker in uttering the complete word, it is found to be more for

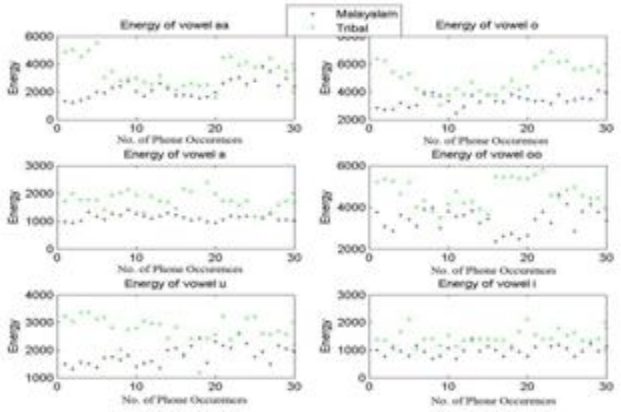

Fig.8. Energy of Malayalam and Tribal speakers in uttering vowels

Tribal than Malayalam speaker. This is observed from the higher amplitudes of speech waveform of Tribal compared to Malayalee. From the figure 9, it can be noted that in uttering the word /vaal/, the amplitude of speech waveform of Malayalam speaker is in the range of 3364 and corresponding amplitude of Tribal waveform is in the range 31820 (which is more).

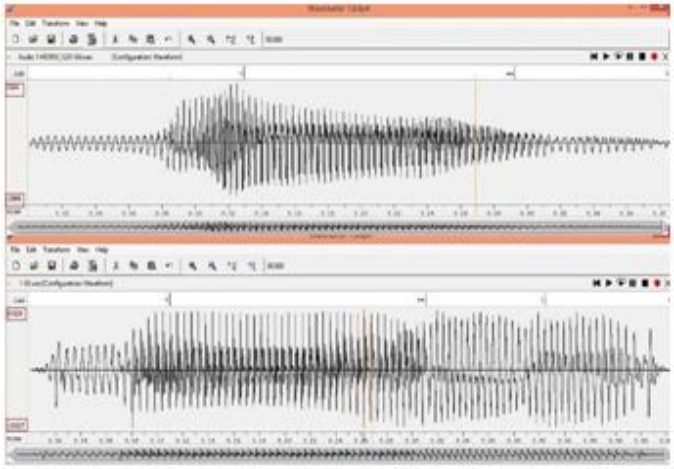

Fig.9. Energy of Malayalam and Tribal word

C. Relation between Phoneme Duration and Energy

It has been observed a direct relationship between the phoneme duration and energy of the speech signal. Plosives and vowel sounds in Tribal speech have a higher duration and energy thancorresponding sounds in Malayalam speech.Nasal sounds have high energy and duration in Malayalam speechcompared to Tribal. This relation can be noted in figure 10 that shows duration-energy plot of a vowel sound.

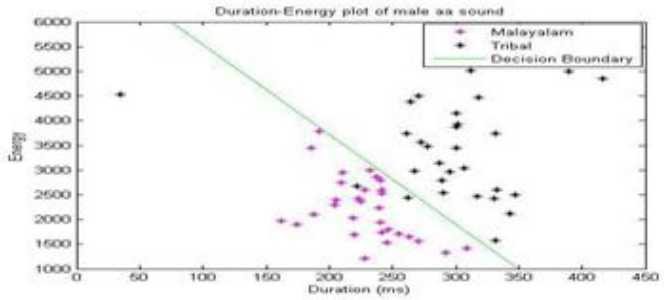

Fig.10.Duration- Energy plot

D. Spectral Flatness

The value of spectral flatness for the pure tone is 0.0 and for white noise 1.0. Spectrum plot appears spiky for pure tone and relatively flat for white noise. It has been observed that the value of the spectral flatness of both Tribal and Malayalam phonemes are in the same range (0.2-0.38) which is near the value for pure tone.

E. Spectral Centroid

Spectral centroid denotes the sharpness of the sound. High value of spectral centroid corresponds to the brighter sound. Spectral centroid of plosive and vowel sounds show a high range for Malayalam speakers compared to Tribals. This means that sound of Malayalam speaker is sharper in uttering plosive and vowel sounds than that of Paniyas. Spectral centroid of nasal and approximant sound shows a high range for Tribal speaker as compared to Malayalam speaker. This gives the inference that sound of Tribal people is brighter when the nasals and approximants are uttered.

F. Pitch

Pitch is the fundamental frequency of a speaker measured in hertz (Hz). Generally it is higher for females than males. It exists only for voiced sounds. So the pitch is absent for unvoiced stops like/p/, /k/, /t/. For all phoneme sounds, pitch of Malayalam females is found to be more than Tribal females.

Mean value of pitch for plosive sounds of Malayalam and Tribal speaker is shown in table II. Pitch of Malayalam malesis higher than Tribal males in uttering plosives.

TABLE II:MEAN VALUE OF PITCH (HZ) FOR PLOSIVE SOUNDS OF MALAYALAM ANDTRIBAL SPEAKER

Phone	Malayalam Male	Tribal Male	Malayalam Female	Tribal Female
/d/	145	134	236	171

Mean value of pitch for nasal sounds of Malayalam and Tribal speaker is shown in table III. Pitch of Tribal males is higher than Malayalam males in uttering nasals.

Mean value of pitch for approximant sounds of Malayalam and Tribal speaker is shown in table IV. Pitch of Malayalam males is higher than Tribal males in uttering approximants.

Mean value of pitch for vowel sounds of Malayalam and Tribal speaker is shown in table V.

TABLE III:MEAN VALUE OF PITCH (HZ) FOR NASAL SOUNDS OF MALAYALAM ANDTRIBAL SPEAKER

Phone	Malayalam Male	Tribal Male	Malayalam Female	Tribal Female
/n/	130	140	240	170
/m/	132	173	221	156
/nn/	141	149	235	180
/mm/	146	155	230	160

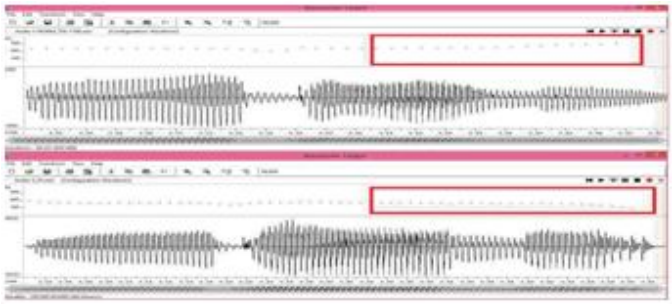

Fig. 11. Pitch contour of Malayalam and Tribal word /ural/

TABLE IV:MEAN VALUE OF PITCH (HZ) FOR APPROXIMANT SOUNDS OF MALAYALAM AND TRIBAL SPEAKER

Phone	Malayalam Male	Tribal Male	Malayalam Female	Tribal Female
/v/	139	130	226	166
/l/	168	160	268	187
/r/	148	142	245	165

Pitch of Tribal males ishigher than Malayalam males in uttering vowels.

TABLE V:MEAN VALUE OF PITCH (HZ) FOR VOWELS OF MALAYALAM AND TRIBALSPEAKER

Phone	Malayalam Male	Tribal Male	Malayalam Female	Tribal Female
/aa/	138	146	240	181
/o/	139	149	252	179
/a/	125	135	246	176
/oo/	143	164	263	186
/i/	122	135	244	183
/u/	135	140	239	173

Pitch Contour: The pitch contour is observed with the help of Wavesurfer tool. Analyzing the pitch contour of Malayalam and Tribal word it has been found that the tribal words follow a decreasing trend. This can be observed from the figure 11.

G. Vowel Triangle

Formant frequency 1 and formant frequency 2 are important in determining the quality of vowel sounds. Units are inhertz (Hz). For assessing the quality of Malayalam and Tribal (Paniya) vowel sounds, their first formant frequencies are plot-ted against second formant frequencies. The formant of vowel /o/ is spread over a wide range for Tribal than Malayalam sounds. Likewise all the vowel sounds taken the formants (F1 and F2) spread over a larger range for Tribal. This means the Tribal (Paniya) phonemes have stronger formants than that of Malayalam speakers, shown in figure 12.

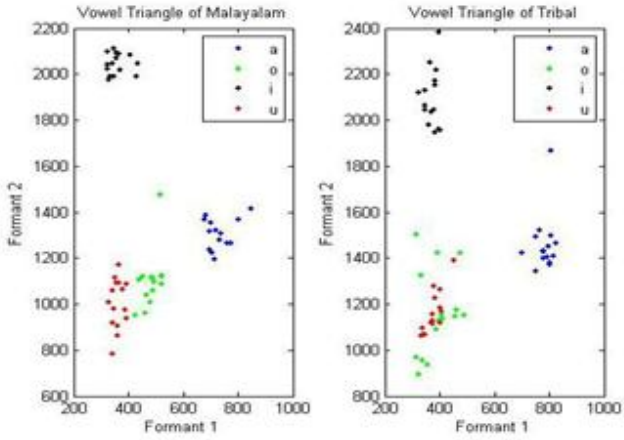

Fig.16.F1-F2 plot of Malayalam and Tribal vowels.

H. Spectral Moment

Spectral moment is a measure to understand the pattern of speech energy present in sound units. Analyzing all the four primary moments-first moment (spectral mean), second mean (spectral variance), third moment (spectral skewness) and fourth (spectral kurtosis), it has been noted that the mean and standarddeviation of the moments of speech sounds follow specific patterns.

The pattern followed by the mean values of S.M. is as follows:

For Vowels and Plosives - Malayalam Male greater than Tribal Male; Malayalam Female greater than Tribal Female. For Nasals and Approximants - Tribal Malegreater than Malayalam Male; Malayalam Female greater than Tribal Female.

The pattern followed by the S.D. values of S.M. is as follows:

For Plosives - Malayalam Male greater than Tribal Male; Malayalam Female greater than Tribal Female. For Vowels - Malayalam Male greater than Tribal Male; Tribal Female greater than Malayalam Female. For Nasals and Approximants - Tribal Male greater than Malayalam Male; Tribal Female greater than Malayalam Female.

V. CONCLUSION AND FUTURE SCOPE

A. Conclusion

The objective of the work was a detailed study between Malayalam and Tribal (Paniya) phonemes. For the purpose various temporal and spectral features like phoneme duration, energy, spectral flatness, spectral centroid, pitch, spectral moments were extracted and analyzed. Analysis was done on the speech corpus collected from the males and females of Malayalam and Tribal language. From the work, it was found that even though the speech signals are highly variable across each speaker in uttering a specific phoneme, when it was analyzed appropriately, distinct patterns were found. Features showed a distinction when analyzed on a phoneme level basis and word level basis. The results suggested that the selected features like phoneme duration, energy, spectral centroid, pitch and spectral moments were effective in discriminating Tribal and Malayalam phonemes.

B. Future Scope

The speech analysis in the area of these two Dravidian languages has a great scope for future work as not much research has been reported. The work can be taken for the design of an automatic speech recognition system and also in classification systems. Instead of using the study on broad phoneme classes the work can be extended for phoneme classes with an extended set of features. This broad phoneme study has a scope for prediction of broad syllable structure and language recognition.

ACKNOWLEDGEMENT

The authors would like to thank Ms. Deekshitha G., Re-search Scholar, Rajiv Gandhi Institute of Technology, Kottayam for the motivation, advice and help to carry out the work discussed in the paper. Also special thanks are due for all the Malayalam and Paniya speakers who generously contributed the speech sample for developing the database.

REFERENCES

[1] Deekshitha G and Leena Mary, “Broad Phoneme Classification Using Signal Based Features”, International Journal on Soft Computing (IJSC) Vol. 5, No. 3, November 2014.

[2] Samuel Thomas, Sriram Ganapathy and Hynek Hermansky, “Spectro-Temporal Features for Automatic Speech Recognition using Linear Pre-diction in Spectral Domain”, European IST Programme Project, 2013.

[3] Samuel Thomas, Kailash Patil, Sriram Ganapathy,Nima Mesgarani, Hynek Hermansky, “A Phoneme Recognition Framework based on Audi-tory Spectro-Temporal Receptive Fields”, In Proceedings of Interspeech (pp. 2458-2461), May 2012.

[4] Siqing Wu, Tiago H. Falk, and Wai-Yip Chan,“Automatic Recognition of Speech Emotion using Long-Term Spectro-Temporal Features”, 16th International Conference on Digital Signal Processing, IEEE, pages1 - 6, July 2009.

[5] J.Bernd T. Meyer, Constantin Spille, Birger Kollmeier, Nelson Morganl, “Hooking up Spectro-Temporal Filters with Auditory-Inspired Representations for Robust Automatic Speech Recognition”, icsi.berkeley.edu/pubs/speech/, May 2016.

[6] Bernd T. Meyer, Suman V. Ravuri, Marc Rene Schadler, Nelson Morgan, “Comparing Different Flavors of Spectro-Temporal Features for ASR”, Interspeech, 12th Annual Conference of the International Speech Com-munication Association, 2011.

[7] Muskan and Naveen Aggarwall, “Analysis of Various Features using Dif-ferent Temporal Derivatives from Speech Signals”, nternational Journal of Computer Applications,(0975 to 8887) Volume 118, No. 8, May 2015.

[8] Douglas O Shaughnessy, “Speech communications-Human and machine”, IEEE press, Newyork, 2nd Edition, 2000.

[9] L Rabiner B H Juang, “Fundementals of Speech Recognition”, Prentice Hall, 2nd Edition, 1993.

[10] Marpe Sora, Jyotismita Talukdar, and P. H. Talukdar, “Formant Fre-quency and Cepstral Method Estimation of Galo Phonemes Using Acoustical Cues”, International Journal of Information and Electronics Engineering, Vol.

3,No.1, January 2013.

[11] Cini Kurian and Kannan Balakrishnan, "Connected digit speech recog-nition system for Malayalam language, Sadhana Vol. 38, Part 6, pp. 1339 to 1346, Indian Academy of Sciences, December 2013.

[12] Kapang Legoh, Utpal Bhattacharjee, T. Tuithung, "Development of Multi- Variability Speech Corpus of Adi Language for Speech Recog-nition Researches", International Journal Of Advanced Research In Computer Science And Software Engineering. Volume 3, Issue 10,Page: 604 to 610, October 2013.

[13] Bong Jik Kim, Son-A Chang, Jing Yang, Seung-Ha Oh, Li Xu, "Relative Contributions of Spectral and Temporal Cues to Korean Phoneme Recog-nition", Korean Linguistics, International Circle of Korean Linguistics. Volume 14, 21-39, 2008.

[14] A textbook about Tribal culture and language, "Nayam". An year long collective action reasearch done in the TTC class in DIET Wayanad.

[15] A textbook of "Kattarum Avarude Kalamozhikalum" Narayanan NAIR (A. R), by Connemara Publication.