IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

Rule-based Model for Automatic Segmentation of Tamil using Supra-Segmental Features

Sakthi Vel S.

Lecturer in Computational Linguistics, Department of Linguistics, University of Kerala, Trivandrum, India

Abstract: This paper investigates the significance of Supra-Segmental features in continuous speech of Tamil. The Supra-Segmental parameters like fundamental frequency, pitch, intensity, time duration, rhythm and etc., can help to signal the syntactic structure of utterances into larger discourse segments and provide additional information for human speech processing. It also provides information about the linguistic structure of speaker's message and the emotional state. In human speech processing, linguistic context and phonological rules help the brain to separate syntactic units into phonemes, syllables, words, sentences and phrases. In the case of agglutinative languages, searching space reduction is very important during automatic speech processing. The entitled study focused on Tamil language, which belongs to the family of languages called Dravidian and noted for its highly agglutinative nature. Here the aim is to examine and prove that prosodic information carried out acoustically by the speech signal can be used to improve the performance of speech processing and to add syntactic, semantic level functionality to it. Therefore, the study introduces a rule-based model which shows the relationship between the Supra-Segmental parameters of phoneme to sentence level and their statistics. Finally, the results were compared to analyse the accuracy and efficiency of the model. The methods introduced here are easily adoptable to other agglutinative languages. Instead of using the prosodic level boundaries the study make use of statistical properties, which are more advanced.

Keywords: Tamil, Supra-Segmental, Fundamental frequency, Duration, Intensity, PRAAT.

I. INTRODUCTION

Supra-Segmental features are an integral part of every spoken language utterance [1]. Pitch, stress, total duration, intensity, fundamental frequency, volume, tempo, pause and etc., are some of the Supra-Segmental features or other ways known as prosodic features. The Prosodic features are those aspects of speech which go beyond phonemes and deal with the auditory qualities of sound. It also encompasses lexical stress, sentential stress and intonation. They are manifested in acoustic cues such as intensity, duration, fundamental frequency and spectral quality. The prosodic features provide insight about the linguistic structure of the speaker's message (the parts of speech of word, syntactic structure of an utterance, etc.), emotional state and etc., Prosody, being an important and integral part of spoken language can be classified as linguistics and emotional prosody.

Linguistic prosody is used to disambiguate and also mark the internal organization of sentence constituents or to convey the intonation contour of a sentence [2]. Prosodic research in recent years have been supported by a number of automatic analysis tools aimed at simplifying the work that intended to study frequency, total duration, intonation and etc. The need to analyse large amounts of data and to inspect phenomena that are often ambiguous and difficult to model makes the prosodic research area an ideal application field for computer based speech processing [3]. The present study uses different types of sentences from Tamil language for the analysis of the relationship between sentence structure and its Supra-Segmental features using sentence and word level segmentation method. The current trends in speech processing systems are to investigate speech in the Supra-Segmental domain. These systems interpret speech as a sequence of phoneme, syllable and etc. and these phoneme sequences are grouped into words, words to phrase, phrase to sentences and sentences to discourse.

The Supra-Segmental system of a language is realized with multiple acoustic features (i.e. intensity, duration, spectral quality and pitch) in a co-ordinated manner [4]. Supra-Segmental features encode rich information structure that helps the listener to locate emphasized words, phrase boundaries speech acts (e.g. statements, questions, continuations, etc.), as well as the speaker's attitudes and emotions [4]. The difference between segmental and Supra-Segmental features appears in the fact that Supra-Segmental features are established by a comparison of items in sequence whereas segmental features can be defined without references to the sequence of segments in which the segment appears, and their presence can be established either by inspection or paradigmatic comparison (i.e. comparison of an item with other items in the phonological inventory). The Supra-Segmental system of language is realized with multiple acoustic features (i.e. intensity, duration, spectral quality and pitch) in a co-ordinated manner [4]. The analysis of Supra-

Segmental features in highly agglutinative languages like Tamil have great significance because these typological categories of languages are characterized with longer word length [7]. The information carried acoustically by the Supra-Segmental features can be exploited in the study of speech recognition. This can improve speech recognition performance and more over this prosodic information can create an acoustical base for the syntactic and semantic processing of human speech and communication. In human speech processing, the word boundaries in continuous speech signal are detected on the basis of Supra-Segmental parameters of the utterance [1].

II. LITERATURE REVIEW

The works done in major world languages have taken into account for the study. The Indian scholars also contributed immensely in the areas of Supra-Segmental features. Among the Indian languages the most noted work in the area of Supra-Segmental is the formulation of word boundary detector for standardized colloquial Bengali and Punjabi (Mandal S. and B. Datta, 2005). But the detector failed to develop as a speech recognizer. The studies have been done in Hungarian and Finnish (both of these languages belong to the Finno-Ugrian language family) also considered as a model for reference. The Supra-Segmental studies in Hungarian shows that an average duration of vowel in Hungarian words is from one to five syllables. In a number of recent works researchers have focused on temporal information for the detection of speech landmarks. The study done on American English by A. Salomon, C.Y. Espy-Wilson and O. Deshmukh, used the Supra-Segmental method as the front end of an HMM-based system for automatic noisy speech recognition. The researchers like Mandal P., and Meloni H., are used multiple cues for detection of phrase boundaries in continuous speech, and integrated these into speech recognition systems.

In 2002, Rahman K.J., and Hossain M.A., developed a word separation algorithm for continuous speech recognition by comparing noise energy and zero crossing with speech of 13 words. The relative importance of intensity, fundamental frequency and duration in the perception of stress have been studied experimentally in several languages, including English (Fry, 1995, 1998; Bolinger, 1958), Polish (Jassem, Morton, 1968), French (Rigault, 1962), Swedish (Westin, Buddenhagen, 1966), and Serbo-Croatian (Rehder, 1968). The renowned scholar Lieberman (1960) studied the acoustic correlates of stress in American English, analysing 25 verb-noun pairs (of the type CONflict -conFLICT), recorded by 16 speakers of American English.

In 2003 J., Hirschberg (Venditti and Hirschberg) summarized the current state of knowledge in intonation and discourse processing available in American English. She described about an intonation discourse interface which can be used in speech technology, mainly for speech synthesis. The method for separating words automatically from a continuous speech in fixed stress languages such as Hungarian and Finnish have been developed by Klara Vicsi and Gyorgy in 1993. These two languages are highly agglutinative, so they are characterized with longer average word length like English. According to Klara Vicsi's study stress is generally on the first syllable of the word, except in the case of conjunctions or articles.

III. PROPOSED METHODOLOGY

The present work makes use of the methodology popularly used by speech technology researchers. Speech is investigated in the Supra-Segmental domain. The acoustic processing of Supra-Segmental information is based on fundamental frequency, intensity and duration. The classification tasks are formulated to model the selected prosodic phenomenon based on Supra-Segmental features (e.g. Stress, intonation), for stress detection or word boundary detection. Here the decision task is interpreted as even detected or not. However, decision can be carried out using more classes like phonological phrase shape and etc.

Decision is based on peak detection and on the obtained classification, and then implements speech segmentation in Supra-Segmental domain, such as word or phonological phrase boundary detection and clause (prosodic phrase) boundary detection. Performance is evaluated by testing, predefined performance measures (recall, precision, recognition correctness and accuracy). The given experiment, examined how word boundaries are detected in continuous speech signal in bound stress languages, like Tamil where the stress is usually on the first syllable of the word. However, it doesn't mean that every word in the language is stressed for instance conjunction, articles etc. But if stress occurs, then it will be in the first syllable [1].

Corpus Creation

The study uses Tamil continuous read speech databases for the examination. The databases were segmented into phoneme, word, phrase and sentence level with marking the boundaries. Most of the Tamil letters have unique sound with their script. But some of the letters have a single script, but sounds differ depending upon on the place they occur. So therefore the letters which change sounds accordance with their place of occurrence must be take into account for

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

the study. For analysing the Supra-Segmental features of Tamil language the data has been collected from 12 native speakers using 6 different sentences forms. The read text of continuous speech is collected from the speakers of both genders (Male and Female) for 5 minutes. The speech data are recoded using the head mounted carbon microphone with a frequency range of 70 Hz- 630 Hz, in a lab environment, sampling rate being 16 KHz. The given table shows the segmentation of speech sounds into sentences, words, syllables and phonemes with extracted values. In order to implement the proposed methodology the present work, has chosen the global speech analyser software PRAAT. The speech analyser software PRAAT is capable of handling multilayer annotations in terms of automatic generation in more efficient way. It is possible because of the scripting language, visualisation, the built-in editors, drawing capabilities, compatibility with external software and many other features of PRAAT. The most important point that needs to be mentioned about PRAAT is that it contains Text Grid format, which can be used or supported in all softwares. The following database includes the value of extract parameters from the selected speech corpus.

Table 1: Speech Corpus and its extracted values

Speech Corpus: Sentence boundary extraction									
Type	Fundamental frequency (F0)/ Pitch (Hz)			Intensity (dB)			Total duration (Sec.)		
	Avg. Male	Avg. Female	Avg.	Avg. Male	Avg. Female	Avg.	Avg. Male	Avg. Female	Avg.
Full Sentence	140.07	186.39	163.23	65.56	63.94	64.75	12.69	12.76	12.7
Speech Corpus: Word boundary extraction									
anta	159.72	200.92	180.32	77.37	84.36	80.86	0.62	0.31	0.46
kuzhanthaikal	191.81	214.16	202.98	79.82	84.57	82.19	0.63	0.57	0.60
vilayaaduvathai	168.80	198.93	183.86	79.95	83.12	81.53	0.76	0.71	0.73
paartherkala	143.97	176.16	160.06	79.22	85.07	82.14	0.77	0.72	0.75
avarkal	158.93	159.31	159.12	80.24	79.60	79.92	0.51	0.52	0.51
Pattaampooccikalaipp	171.95	188.94	180.44	76.24	81.10	78.67	1.06	1.10	1.08
pontru	102.88	111.85	107.36	71.41	79.86	75.63	0.41	0.34	0.37
angum	150.93	178.21	164.57	79.17	78.71	78.94	0.48	0.48	0.48
ingum	129.79	150.34	140.06	79.80	81.6	80.70	0.43	0.39	0.41
aadi	128.89	137.31	133.10	81.45	84.65	83.05	0.38	0.78	0.58
thirivathai	136.91	165.84	151.37	79.35	83.06	81.20	0.38	0.78	0.58
paarppatharku	122.63	144.65	133.64	75.81	83.98	79.89	0.76	0.78	0.77
yethanai	160.64	178.81	169.72	45.92	63.98	54.95	0.51	0.52	0.52
azhagu	109.87	117.64	113.75	79.78	74.12	76.95	0.44	0.41	0.42
kuzhanthaikal	196.30	200.27	198.28	80.58	84.12	82.35	0.61	0.54	0.58
vilayaadithaan	123.93	177.95	150.94	73.80	83.41	78.60	0.82	0.90	0.86
valaravendum	105.54	168.97	137.25	82.11	76.65	79.38	0.71	0.65	0.68

Word Boundary detection based on stress level

For stress detection fundamental frequency (Hz) and energy level (dB) in the middle of syllables were measured. The Autocorrelation method was used for the determination of fundamental frequency, the autocorrelation function for **x (n)** discrete signal is:

$$R(k) = \sum_{k=N-n-1}^{N} x(n)\, x(n+k) \quad \textbf{(1)}$$

$\mathbf{F_0}$ at the $\mathbf{i^{th}}$ frame $\mathbf{F_0}$ **(i)** was obtained after median filtering:

$$F_0(i) = \boldsymbol{med}\,\{\, F_0(i-3)\,,\, F_0(i-2)\,,\, F_0(i-1)\,,\, F_0(i)\,,\, F_0(i+1)\,,\, F_0(i+2)\,,\, F_0(i+3)\,\} \quad \textbf{(2)}$$

The energy **E (i)** was calculated with an integration time of 100 ms:

$$E(i) = \frac{1}{M} \sum_{n=i-\frac{M}{2}}^{i+\frac{M}{2}} x^2(n) \quad \textbf{(3)}$$

Here (**M)** is the number of samples pro 100 ms.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

Stress is interpreted as an emphasis on a syllable of a given word. A language is called fixed stress language, if stress is always bound to a given syllable of a word. Tamil is a stress language and the stress usually occurs on the first syllable of the word stressed. The other fixed stress languages are Malayalam, Finnish and Hungarian etc. And it is commonly seen that in highly agglutinative languages, almost all the words have some stress normally on the first syllable except in case of conjunctions and articles. Any prosodic features can govern stress in a speech, like the rise in the fundamental frequency, an augmented intensity of the syllable, duration lengthening etc. All the above mentioned features may provoke stress sensation depending upon the language in which the feature or feature combination is used to express stress. In Tamil, a considerable number of word boundaries can be identified based on stress detection. And stress detection is carried out on the bases of its prosodic attributes, fundamental frequency and intensity (energy). Using this approach, word boundaries followed by a stressed word can be detected. The prosodic features can be measured of each vowel in each syllable, or continuously on the speech signal.

Boundary detection of the Word Unit

If the stressed syllable can be detected, the word boundary can be marked by finding the minimum just before the stressed point. Hence the stressed vowel may be situated in the onset or even in the nucleus of the syllable, while searching the minimum point, the absolute minimum within an interval just before the stress can be found. The length of this interval is typically a half-syllable in order to avoid overlap with previous word. The present works concentrate only on rule-based approach.

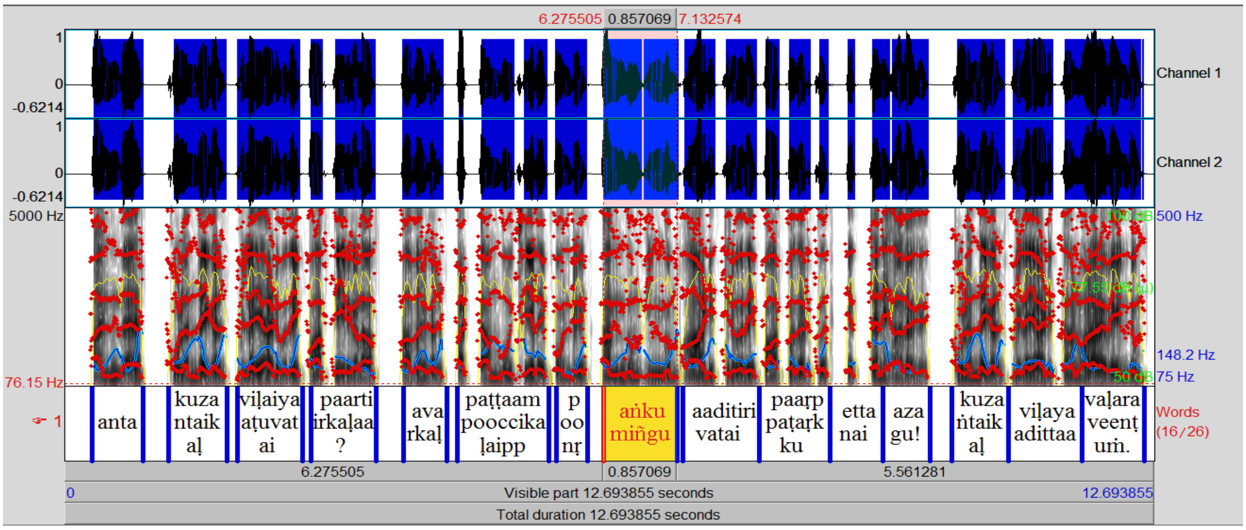

Figure 1: boundary detection based on word unit

Peak detection algorithm

The present work developed a peak detection algorithm, which observes the selected prosodic features over a sliding window. The peak detection algorithm, which **(M)** and variance **(Ϭ)** of fundamental frequency and/or energy measures are calculated over the sliding window, and then a threshold value is calculated as follows:

$$\mathbf{K = M + K * \sigma} \qquad \mathbf{(4)}$$

Where (**K**) is a constant value, the interval of approx is **0.5-1.5**. Prosodic features in the middle of the window are controlled whether they override the threshold or not. If so, a peak is found and stress is detected on the corresponding syllable. This stressed position can be mapped to a word boundary (based again on energy or fundamental frequency constraints). The threshold used to control stress on $\mathbf{i^{th}}$ syllable is:

$$K_i = M(x_{i-A}, x_{i-A-1}, \dots, x_i) + c * \sigma(x_{i-A}, x_{i-A-1}, \dots, x_i) \quad if\ i > A \qquad (5)$$

$$K_i = M(x_1, x_2, \dots, x_A) + c * \sigma(x_1, x_2, \dots, x_A) \quad otherwise \qquad (6)$$

Where **(A)** is the length of the sliding window measured in the number of syllables. The peak detection algorithm was performed on **F0** and energy level data. The data streams obtained by computing absolute value of **F0** and energy level differences between two neighbouring syllables as follows:

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

$$M_i = \frac{1}{A}\sum_{j=i-A}^{i} |\Delta x_j| \qquad (7)$$

$$\sigma_i^2 = \frac{1}{A}\sum_{j=i-A}^{i} (M_j - |\Delta x_j|)^2 \qquad (8)$$

Again we used a sliding window **(5, 6)** to compensate for decreasing dynamic range of speech towards the end of the sentence. Instead of the pure fundamental frequency or energy measures, their differentials can also be used. Differences of these features can be calculated from syllable to syllable or interpreted as a derivative of a continuous contour. After taking the absolute value of the difference, peak detection is carried out. In this context using a sliding window is of great advantage, as it allows adapting to varying dynamic range.

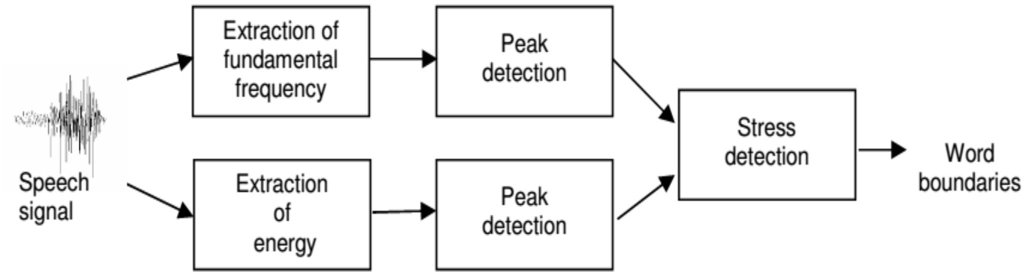

Figure 2: Stress and word-boundary detector system

Evaluation

For speech recognition tasks, the accuracy is more critical than effectiveness, the present work expects that if a word boundary is predicted it should be accurate at least around 70 %. It is true that, higher the effectiveness the more robust the system will be, but we cannot allow this at the expense of falling accuracy. To evaluate the results the obtained prosodic segmentation is compared with the original one. Two measures are used to present the results. First one is the **correctness** which denotes whether a unit boundary is predicted by the investigator developed algorithm was correct or not (in %).

$$Corr\ [\%] = \frac{\#\_of\_correctly\_marked\_word\_boundaries}{\#\_of\_all\_marked\_word\_boundaries} * 100$$

The second measure is the **effectiveness** which says how many word boundaries were found at all (in %): The second measure is expected to be more than 70 %, since not all the words in the speech emphasized. Articles and conjunctions can't be separated in this way because they are not emphasized. As for the speech recognition tasks we are concerned that whenever a word boundary is predicted it should be detected correctly. The correctness is more critical than effectiveness.

$$Eff[\%] = \frac{\#\_of\_correctly\_marked\_word\_boundaries}{\#\_of\_all\_real\_word\_boundaries\_(in\_reference\_transcription)} * 100$$

Rule-based approach

Using the peak detection algorithm for stress localization, the researcher investigated the performance in function of **(C)** constant and (**A)** sliding window length parameters. The below Table 2: shows the results in six columns depending upon which combination of prosodic parameters **(only F_0, only Energy, F_0+Energy, only ΔF_0, only ΔEnergy, only ΔF_0+ ΔEnergy)** stress detected, as expected, rising **(C)** constant results in a higher accuracy with a more considerable fall in effectiveness. The more accurate results were obtained by detecting stress on the basis of fundamental frequency and energy level changes (**ΔF_0**) from syllable to syllable.

Table 2: Accuracy and effectiveness of stress detection values

A [#of syllables]	C	Accuracy/Effectiveness [%/ %]					
		F_0	E	F_0+E	ΔF_0	ΔE	ΔF_0+ΔE
9	0.5	49/41	46/29	95/70	76/24	59/21	135/45
10	0.9	52/32	46/23	98/55	78/21	61/17	139/38
12	0.11	52/27	45/20	97/47	79/19	62/15	141/34

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified

Vol. 6, Issue 9, September 2017

Results

The present system developed by the investigator is convenient for automatic segmentation of word units. An example is given in Figure 3: of how the developed segmentation method works on word level. Time function of the speech signal, **F0** and energy level contour are visible on the screen. The following spectrogram shows the audio-visual segmentation by the automatic prosodic segmentation. Segmentation accuracy means the correctness in the determination of word unit.

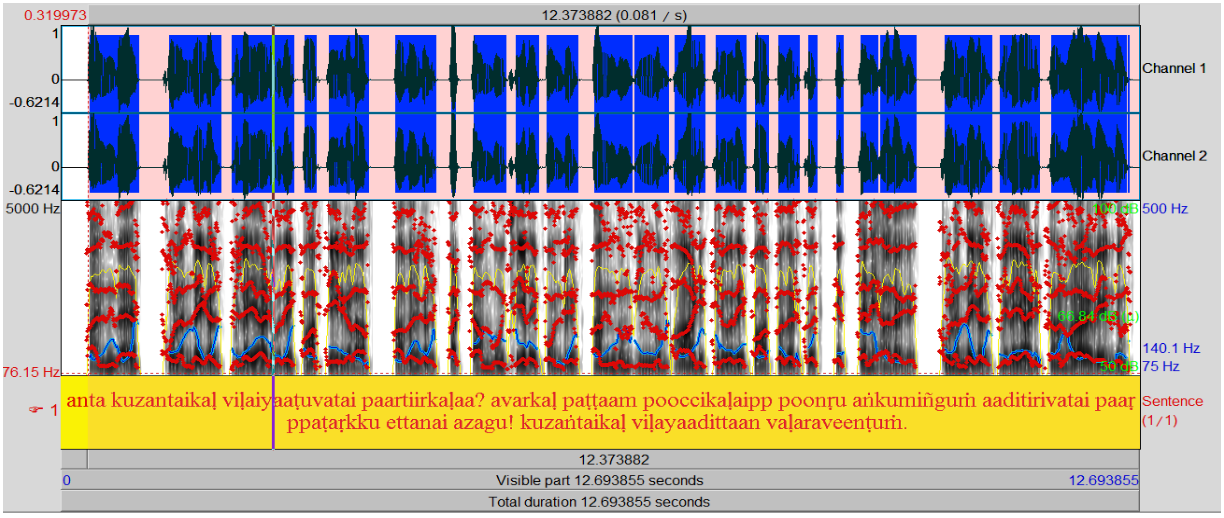

Figure 3: Sentence *'anta kuzantaikaḷ viḷaiyaaṭuvatai paartiirkaḷaa? avarkaḷ paṭṭaampooccikaḷaipp poonṛu aṅkum iñguṁ aadi tirivatai paaṛppaṭaṛkku ettanai azagu! kuzaṅtaikaḷ viḷayaadittaan vaḷaraveenṭuṁ'*

According to proposed methodology accuracy and effectiveness of stress detection based on peak detection algorithms for Tamil data in function of sliding window width (**A)** and of (**C)** constant present in threshold calculation for six prosodic data patterns. Table 3: The following table, shows that correctness and effectiveness of boundary determination of word unit with Peak detection for Tamil language.

Table 3: correctness and effectiveness of boundary determination

Used prosodic parameters	Language	Training corpus	Correctness/ Effectiveness [%/ %]
$F_0 + \Delta F_0 + \Delta^2 F_0 + E + \Delta E + \Delta^2 E$	Tamil	12 persons (6 Male + 6 Female)	63.89 /70.39 (approximately)

IV. SCOPE OF THE STUDY

The present study ensures that the integration of a Prosodic Recognizer into an Automatic Speech Recognition (ASR) system can help to reduce the search space and thus improve speech recognition performance. The importance of this search space reduction is greater in the speech recognition of agglutinative languages such as Tamil, Malayalam, etc. The present study intake a special challenge to implement a prosodic framework that ensures the acoustic level extraction and processing of Supra-Segmental speech attributes. This is intended to be used for the improvement of Automatic Speech Recognition classification tasks which allow syntactical processing and segmentation (segmentation for phonological phrases, clauses, sentences) and higher level semantic processing of speech message (modality recognition). This framework for speech recognition can also be used as a basis that provides acoustically processed Supra-Segmental information to be used in Natural Language Processing (NLP) tasks. The present work can be used in computer-aided speech training systems for acquiring the use of stress and intonation in speech [7].

V. PROBLEMS AND LIMITATIONS OF STUDY

Using Supra-Segmental features in automatic speech recognition in order to increase its robustness is a commonly viewed tendency now days. Some trials were conducted in the mid-eighties. But to exploit such knowledge in automatic speech recognition system in its full fledged form is a great failure in this field. The study on Automatic Segmentation of continuous speech using Supra-Segmental features has some difficulties too:

- *Significant contextual variability of prosodic knowledge*: (type of speech, speaker, structure and content of sentences, nature of the environment, etc.).
- *Complexity of relations*: between prosodic information and various linguistic organization levels of a message.
- *High variability in acoustic features:* Supra-Segmental phenomena are manifested as their acoustic correlates in speech, but acoustic signal is modulated by many factors, including individual speaker variations, lexical and sentential contexts, communicative semantics, non-linguistic factors such as emotion, etc.
- *Lack of consciousness about communicative significance of Supra-Segmental features*: Supra-Segmental features can be used to express pragmatic meanings which are related to the speaker or the discourse, e.g. a speaker may change their pitch to express disinterest in the speech.
- Problems encountered with accurate measurement of prosodic parameters, and their possible integration on a perceptual level [5].

Along with the above mentioned difficulties, the Automatic Segmentation of continuous speech using prosodic features has some limitations also. This is more crucial for Tamil language, because of its high agglutinative nature, which hinders the adoption of standard methods of speech recognition developed for other typological category of languages.

The major limitations are:
- There are sentence types influenced primarily by intonation. The present study does not consider declarative, interrogative, exclamatory and imperative sentences.
- Complexity of relations between prosodic information and various linguistic organization levels of a message.

VI. APPLICATIONS OF THE STUDY

The human language processing takes place in brain and it is the linguistic content and phonological rule that helps the brain separate language and its syntactic units [7]. Natural language processing (NLP) is concerned with how system process human language. In NLP, speech recognition can be used as a basis for syntactic and semantic analysis and parsing (hierarchical weighting of information, deeper sentence analysis in automatic translation systems, etc.) or voice mining, voice enable search engine, speech document classifications and etc. The experiment result can be used in automatic speech recognition systems and highly useful in linguistic research, especially for the research related to spontaneous speech. The Rule based segmentation allows the automatic placement of punctuation marks in ASR output. The findings are also applicable for natural language processing for syntactic and semantic level related tasks such as topic-comment classification, sentence- clause level segmentation of speech, focus detection, sentence analysis, and modality classification. They allow an accurate detection of prosodic phrases.

VII. CONCLUSION

The present work on word level prosodic segmentation method based on measuring fundamental frequency and energy level functions provided promising results. The word boundaries can be marked with acceptable correctness, even if we are not able to find all of them. Two measurements such as correctness and effectiveness were used to describe the behaviour of the prosodic segmentation system. The Word boundaries are detected with acceptable correctness and effectiveness in agglutinative languages like Tamil, Malayalam, etc.

Moreover, these results ensure that the integration of a prosodic recognizer into a ASR system can help to reduce the searching space and thus improve speech recognition performance. The importance of this searching space reduction is great in the speech recognition of agglutinative languages such as Tamil, Malayalam and etc., whereas the possible number of word forms is more than hundreds of millions. The methods introduced here are easily adaptable to other agglutinative languages. Instead of using the prosodic level boundaries the statistical properties are used which is more advanced. Summarizing the results of present experiment it becomes vivid that it is worth continuing research in this field. In fact further investigations are needed in this domain of language and the researcher hopes that present study may be helpful for triggering further research in this area of language.

REFERENCES

[1] Klara Vicsi- Gyorgy Szaszak. Automatic Segmentation of Continuous Speech on Word and Phrase Level based on Suprasegmental Features. Forum Acusticum 2005 Budapest. Pp. 2669-2673.
[2] Mili Mary Mathew and Jayashree. S. Bhat. Aspects of Emotional Prosody in Malayalam and Hindi. The Buckingham Journal of Language and Linguistics Volume 3. Pp. 25-34.
[3] Antonio Origlia, Iolanda Alfano. Prosomarker: a prosodic analysis tool based on optimal pitch stylization and automatic syllabification. Pp. 997-1002.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

[4] Helen Meng, Chiu-yu Tseng. Studying L2 Suprasegmental Features in Asian Englishes: A Position Paper. The Chinese University of Hong Kong.
[5] Klara Vicsi, Gyorgy Szaszak. Automatic Segmentation of Continuous Speech on Word Level based on Suprasegmental Features. International Journal of Speech Technology 2005. Pp. 363-370.
[6] Gyorgy Szaszak, Klara Vicsi. Using prosody for the improvements of automatic speech recognition. Info communications journal. Volume LXIII. 2008/7. Pp.35-40.
[7] Sakthi vel S. Rule- based model for Sentence level Automatic segmentation of continuous speech of Tamil using Supra-segmental features. Unpublished M.Phil. Computational Linguistics Dissertation. Pp. 7-8, 11.
[8] M.A. Anusuya and S.K. Katti. Speech Recognition by Machine: A Review. International Journal of Computer Science and Information Security (IJCSIS), Vol. 6, No. 3, 2009. ISSN 194-5500. Pp.181-200.
[9] Directorate of Distance Education, Annamalai University (2005). P.G. Diploma in NLP. Computer Applications to Language Studies. Annamalai University Publication.
[10] Directorate of Distance Education, Annamalai University (2005). MA Linguistics. Phonetics and Phonology. Annamalai University Publication.
[11] Directorate of Distance Education, Annamalai University (2005). MA Linguistics. Introduction to Linguistics. Annamalai University Publication.

BIOGRAPHY

Sakthi Vel S., was born on February 13, 1989. He received the B.SC., MCA, and M.Phil. (Computational Linguistics) Degree from University of Kerala, Kerala and MA (Linguistics) from Annamalai University, Tamil Nadu. He is working as Lecturer in Computational Linguistics in Department of Linguistics, University of Kerala, India since 2014. His research interests include Computational Linguistics, Speech processing, Tamil computing.