The 6th Intl. Workshop on Spoken Language Technologies for Under-Resourced Languages
29-31 August 2018, Gurugram, India

Prosodic Analysis of Non-Native South Indian English Speech

G Radha Krishna[1], *R Krishnan*[2], *V K Mittal*[3]

[1]Associate Professor, VNRVJIET, Hyderabad
[2]Adjunct Faculty, Amritha University, Coimbattore
[3]CEO, Ritwik Software Technologies, Hyderabad

gunturrkr@gmail.com, drrkdrrk@gmail.com, drvinaykrmittal@gmail.com

Abstract

Investigations on linguistic prosody related to non-native English speech by South Indians were carried out using a database specifically meant for this study. Prosodic differences between native and non-native speech samples of regional language groups: Kannada, Tamil, and Telugu were evaluated and compared. This information is useful in applications such as Native language identification. It is observed that the mean value of pitch, and the general variation of pitch contour is higher in the case of non-native English speech by all the three groups of speakers, indicating accommodation of speaking manner. This study finds that dynamic variation of pitch is the least for English speech by native Kannada language speakers. The increase in standard deviation of pitch contour for non-native English speech by Kannada speakers is much less at about 3.7% on an average. In the case of Tamil and Telugu native speakers, it is 9.5%, and 27% respectively.

Index Terms: Pitch, regional accent, prosody, Indian English, pitch contour, standard deviation.

1. Introduction

Perceptually salient cues matching human capabilities, by which different language groups can be identified, are useful in categorizing speakers [1]. Humans use prosodic information additionally, to judge accents, which also helps them to differentiate languages easily [2]. In a spoken language, prosody describes the patterns of stress, intonation, and fluency. Accent of a particular language group has distinct prosody. Addition of this knowledge would improve automatic speaker [3]. Non-native speech deviate from native speech basically in prosodic features: *Fluency*, *word stress* and *intonation*. Prosody is linked to the syllables of a language, which deviate in terms of *melody*, *timing* and *stess*. The long term speaker-specific characteristics such as *pitch gestures, accents* and *stress* in a spoken utterance mainly constitute the Prosodic cues [4]. Apart from vocabulary and grammar, acquisition of prosody, which vary among languages is important for language learnerning [5]. Analysis and classification of utterances that belong to specific groups of learners is the main objective of Native Language Identification (NLI) [6]. Dynamics of F_0 reflecting certain learned traits is important in technological application of accent characteristics for forensics and security [7, 8]. To understand the role of intonation and syllable stress in the foreign accent, dynamics of pitch contour across three different language groups in South India: *Kannada* (KAN), *Tamil* (TAM), and *Telugu* (TEL) were studied.

Previous studies on the differences between native and non-native speech provided clear separation of *fluency* and *duration* related features (duration of phones and speech rate). The variations of *intonation* (pitch contour) were investigated mostly for a set of two languages [9]. The study by [10] observed that the speech style affects the *pitch range*, which is a linguistically relevant aspect of pitch. Prosodic or suprasegmental features involve pitch and intensity of the speech signal, and are controlled by the laryngeal activity [11]. These are very important issues when learning a foreign language [12]. Prosodic contours were used to evaluate the expressiveness of the read speech produced by both native and non-native speakers [13, 14]. These variations correspond to pitch, intensity and length at the perceptual level. Languages have characteristic sound patterns, *pitch contours* are widely used for decoding this acoustic signature as in *language identification*(LID) [15]. The source at acoustic level approximately represents prosody at the linguistic level. The levels of separation between native and non-native speech can be in terms of *standard deviation* (SD) of F_0 contour. The non-native corpora exhibits greater F_0 SD and higher overall pitch quantiles [16]. A hierarchy of speech characteristics, related speaker traits, and possible speech features are listed in Table 1.

Producing correct prosody to achieve a natural and fluent pronunciation when compared with native speaker is a very relevant issue in Computer Aided Language Learning (CALL) [5, 13, 17]. In the context of prosody, *intonation* is more used than *energy* and *duration* features. Native traits located at word and sentence levels help listeners structure the speech signal. These suprasegmental parameters can be used to model non-native English prosody [18]. Text-Independent methods can also be employed to study the degree of nativeness, by analysing the pitch contour. To estimate F_0 correctly and compare among native and non-native languages, a model that does not incorporate any phonology of the specific languages is desired [19]. The problem of pronunciation quality evaluation in non-native English speech from intonation point of view is aimed in this study.

English is widely used second language in India and elsewhere in the world [20]. Many users of services such as education, security and forensics use Englsih non-natively. Non-native prosodic traits limit proficiency in a second language (L_2), and also the mutual understanding. Identification of speakers, classification of their dialectal zones are improtant tasks in a multilingual country like India [21, 22]. Developing an accent classification algorithm that can extract the necessary information to identify the linguistic zones is the main objective of present research. The F_0 patterns were examined in the three regional varieties of English speech to bring out information pertaining to the identification of native language. The goal is to evaluate the use of suitable features of non-native prosody, which help in predicting the the type of foreign accent from text-independent South Indian English speech. The paper is organized as follows: Section 2 presents the details of participants and material in speech data. Section 3 describes the techniques available for assessment of linguistic prosody, and

10.21437/SLTU.2018-15

Table 1: *Summary of speaker traits and related speech features. [23]*

Speech characteristic	Speaker trait	Speech feature
Lexical, Syntactic (Idiolect,Semantics, Pronunciations, dictions, Idiosyncrasies)	Socio economic Educational status (Language use and sentence construction)	Vocabulary, Word arrangement & grammatical cues.
Prosodic (Rhythm,Intonation, articulation rate etc.)	Personality type, Parental influences	Durational features. Pitch dynamics, Energy (likely to be Text/time dependent).
Low level acoustic features	Anatomical structure of speaker's vocal apparatus	Short-time spectrum, Predictor coefficients, Intensity, Pitch.

Section 4 presents a procedure for prosodic characterization of accent. Experiments that were carried out and results obtained are given in section 5, and the key observations are discussed in section 6. Conclusions drawn are presented in Section 7.

2. Database

Manifestation of foreign accent as distinct patterns of articulatory behavior in the multilingual Indian scenario is considered in this study. Using high-level information to further enhance short-time, cepstral based systems has been popular in mainstream speaker recognition technology [6]. Current study presents a case when large resources of spoken English utterances to train speaker identification system parameters are not available [24]. This corpus is developed to explore the linguistic features that constitute the prosodic typology. Native speakers of *Telugu (ISO 639-3 tel), Tamil (ISO 639-3 tam)*, and *Kannada (ISO 639-3 kan)* produced comparable utterances in English and in their native languages. These languages are divergent and differ in the amount of Sanskrit they contain. Indian languages are mostly syllable-timed, whereas English is stress-timed. Studies on melody metric indicates Telugu is syllable timed whereas Tamil is Mora timed. [4, 25]

2.1. Participants

Totally 120 speakers participated in the study. The speech material comprises of South Indian speakers speaking in one of the three native languages or in non-native English. The speakers are separated into two sets. The first set consisted of speakers with only one recording, speaking either in one of the native languages mentioned above, or in English. The second set consisted of recordings of speakers speaking in both native language and in English. The native speech data base consisted of recordings of 20 native speakers in each of the three languages KAN, TAM, and TEL. A total of 6000 seconds of speech database is developed for each of the above three native languages. These details are shown in Table 2. A different subset of speakers, from each the above mentioned language groups, who are capable of speaking English is selected as the non-native speech set. This set of data base contained of English utterance of the speakers belonging to one of the above three languages as *mother tongue*. English test sample for a duration of 60 seconds is collected from 25 speakers in each native language group. The details of these non-native English speakers are presented in Table 3. The non-native English utterances are recorded under identical conditions as native speech samples. English speech samples are collected from a set of speakers with wide ranging geographical distribution within a state. The speakers in the native set are considered representative of the regional languages KAN, TAM and TEL. Compatriots of these speakers who are more exposed to reading, listening and even talking English are considered for present analysis.

Table 2: *Distribution of native speech set.*

Native language	No. of speakers Female	Male	Duration(sec)
KAN	10	10	6000
TAM	10	10	6000
TEL	10	10	6000

Table 3: *Distribution of non- native speech set.*

Native language	No. of speakers Female	Male	Duration(sec)
KAN	10	10	60
TAM	10	10	60
TEL	10	10	60

2.2. Speech Material and procedure

Recordings of speakers were made in quiet office room conditions using Logitech h110 microphone. It was ensured that data is as balanced as possible. Recordings of the speech are made with text material from general topics, mostly on Personality development, and technical literature, with speakers reading aloud under relaxed conditions. Gender weightages are almost equally distributed in both the native and non-native data sets. Age groups are according to SPEECON convention [26], with a bias of the younger age groups. The speakers in the native set are so chosen that they are not from border places of neighbouring states, thus avoiding dialectal variation. Most of the speakers in the non-native English speech set are closely associated with academic institutes. Speakers removed from the consideration of the non-native speech set are the English teaching faculty, and those who are convent educated.

3. Assessment of Linguistic Prosody

Prosodic structure is a critical aspect of language contact, and gives important information related to the speaking habit of a person [4]. Prosodic phenomena located on word level and above, help listeners to structure the speech signal and to process the linguistic content successfully. Empirical studies substantiated that hardly any second language learner after *critical period* manage to acquire consistently native pronunciation. The characteristic speaking style of an individual is influenced by his/her native language [27]. Non-native speake of English is likely to use phonemic inventory, and prosodic patterns of her/his native language [15]. The English prosody of Kannada, Tamil, and Telugu speakers can sound strange as they do not display the native sturcture of Englsih. An attempt is made in the present study to capture prosodic idiosyncrasies of speakers belonging to different native language groups [28].

Table 4: *Major Text Independent features used in prosodic analysis.*

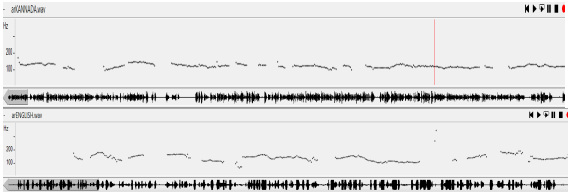
Prosodic features	Factors that influences speech
Dynamics of F_0 contour	Identity of sound unit, its position with respect to phrase or word; Speaking style of the speaker; Intonation rules of the language; Type of sentence (Interrogative, Declarative)
Intonation, Rhythm, Stress	Attitudinal, Accentual, Discourse, Grammatical

3.1. Acoustic Analysis of Pitch Contours within and across Linguistic groups

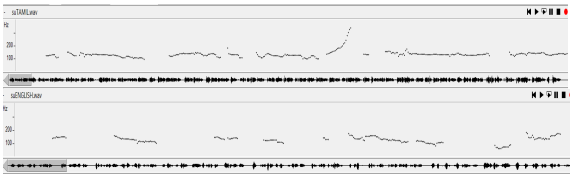
Linguistic prosody is manifested in terms of changes in *fundamental frequency (F_0), duration* and *energy* [4]. Prosody probably plays an auditory role similar to that played by facial expression and body language in the visual world. Prosody between languages varies greatly and *pitch contour* is one of the prosodic cues that can be extracted directly from speech signal. Pronunciation quality evaluation from intonation point of view in second language learning is yet to be explored fully [5, 17]. The melodic features of regional accent is an outcome of both spectral changes and pitch variation. Use of prosodic information to enhance state-of-art acoustic systems has been very popular [29]. In addition to intonation (*pitch contour*), which is reflected in the change of F_0 in utterances, prosody encompasses patterns of phoneme and pausing durations [30]. Table 4 shows some of the features useful for detecting non-native speech without annotation of prosodic events. Prosodic statistics are usually obtained by performing different measurements of *energy, duration* and *pitch*, which are derived supra segmentally [32]. The *rate of speech* and *pause* patterns are other available cues that would provide valuable information regarding speaker class and condition. Representation of the pitch contour with suitable features can bring out the language-specific information present in it.

3.2. Imitation aspect of Language Learning

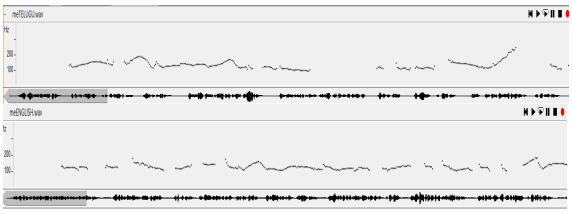
Recent technologies require understanding of the characteristics of speech to create systems where machine dialogues to be tailored to particular group of clients. The background of speaker plays an important role when learning a foreign language [33]. Learners having different native languages belonging to related linguistic backgrounds have similar developmental processes when acquiring a second language. In the process of learning both native and foreign languages, *imitation* is used. It is also used for accommodation of speaking manner in a community [34]. Most clear process of imitation in language acquisition is phonetic and phonologival reproduction. The learned speaking habits are usually influenced by the social environment and also by native language in the *critical period* of learning, which identify where a person is from regionally [35]. Dynamics of F_0 contour related to a sound unit are affected by factors such as identity of sound unit, its position corresponding to the phrase/ word , its context, the speaking style of the particular speaker, intonation rules of the language, type of the sentence etc. [36]. Apart from non-deliberate factors mentioned above, deliberate modifications vary across speakers. Voice imitation provides an opportunity to investigate aspects of human voice needed to be altered to deceive a listener.



(a) Native Kannada speker speaking *KANNADA* (top), *ENGLISH* (bottom)



(b) Native Tamil speker speaking *TAMIL* (top), *ENGLISH* (bottom)



(c) Native Telugu speker speaking *TELUGU* (top), *ENGLISH* (bottom)

Figure 1: *Typical F_0 contours of Native speech and Non-native English of (a) Kannada, (b) Tamil, and (c)Telugu native speakers.(using Wave Surfer tool [37])*

4. Intonational differences across Native and Non-native speech

Intonation is one of the most certain features that makes it possible for a native speaker in recognizing whether an utterance has been produced by a native speaker or not [17]. Intonation helps in identifying the grammatical structure in speech, similar to punctuation in writing. Listeners can discern a speaker's regional accent from hearing the intonation alone in filtered speech [38, 39]. Linguistic features constituting a prosodic typology include lexical accent, intonation and rhythm. To convey a wide variety of meanings or situations, within the framework of grammar, intonation is used. Understaning the role of L_1 typology in L_2 intonation needs investigations [8]. Prosody describes the patterns of stress, intonation and pausing in a spoken language, intonation is reflected in the change of fundamental frequency (F_0)in utterances. The differences in the physical structure of the vocal folds among speakers is manifested in the variation of the distribution of F_0 values. Different speaking style and accent among different speakers is reflected in differences in dynamics of F_0 contour and energy contour. When the same text is repeated by a given speaker, the characteristics of F_0 contour are consistent, and this is true across speakers.Pitch movement is an important feature of the melodic component of the non-native speech [5]. In most of the South Indian cities, at least three languages are spoken: the official language of that state, English and the one of the languages of neighbour-

Table 5: *Typical statistics of F_0 contour for native and non-native speech for single male and female speaker from each native language group Kannada, Tamil, and Telugu.*

Language Group	SD when speaking native language		SD when speaking English	
	MALE	FEMALE	MALE	FEMALE
KANNADA	32.2	30.7	32.3	31.9
TAMIL	21.7	21.8	28.9	25.2
TELUGU	21.5	35.5	22.9	41.2

Table 6: *Percentage increase in Standard Deviation of pitch contour from native language speech to English speech (using two non-overlapping sets of 20 speakers from each native language group Kannada, Tamil, and Telugu.*

Language group	Male	Female
Kannada	0.9	6.5
Tamil	9	10
Telugu	33	21

ing states. Studies were carried out to ascertain the way the linguistic background influences the intonational attributes of fundamental voice frequency F_0 across three regional varieties of English [31]. In many approaches that apply prosody to either Language Identification (LID) or Speaker Recognition, extracted features are based on statistics of pitch/energy contour segments or piecewise linear stylization of pitch/energy contours [15]. While dealing with prosody in the second language learning, *intonation* is the main consideration than *energy* and *duration*. With major emphasis on application of NLI to the learner's non-native speech, the focus is on research of (L_2) prosodics. These include *rhythm* and *intonation*, which may imply speech acts like making a statement [20]. Present work helps in development of automatic intonation classification for (L_2) speech.

5. Experiments and Results

A comparative study of voice produced by bilingual speakers, who are natives of South India was made for exploring the acoustic characteristics of native and non-native speech samples. The role of accent in voice identification is investigated as explained below. The speakers in this study were asked to speak in their mother tongue or in English, and in few cased in both languages. F_0 data was extracted using Wave Surfer software [37] for each of these speakers. The focus was mainly on pitch, since it is one of the most important characteristics of prosody and helps in predicting human intonation rating. The pitch frequency was extracted using the "pitch contour" function of the *Wave Surfer* and F_0 data was extracted for all the speakers whose back ground is described in section 2. Figures 1(a), (b), and (c) shows the pitch frequency contours of native and foreign accents in the speech data described in section 2. The difference in F_0 contour between native and non-native speech for a single male speaker from each group cab be clearly seen from these figures.

Experiments were also conducted on native and non-native speech samples of bilingual and multilingual speakers. Twenty exemplars were analysed from each of the language groups *KAN*, *TAM*, and *TEL*. All the speakers were asked to speak in their mother tongue, English and in some cases in other Indian language spoken in the neighbouring states in which they have the working knowledge. Table 5 shows typical standard deviations (SD)of pitch frequencies for speech data of native speech uttered by a male speaker in each language group, and that of the same speakers in their non-native English speech. It is evident from Table 5 that the SD is higher in non-native speech in all the three language groups for both male and female. It can be observed from deviation scores presented in Table 6, that in general, native Kannada language speakers when speaking English, has smaller variation in F_0 values compared to native Tamil or Telugu speakers, in spite of the variation in speaker characteristics.

6. Discussion on Prosodic Analysis of South Indian Languages

As expected from previous studies on non-native English production, the non-native read speech showed greater F_0 standard deviations and higher overall pitch.

- This study explored the main question of evaluating the prosodic competence by studying the accent patterns of non-native English speech produced by three native language groups in South India.
- The results obtained clearly indicate that the mean value of non-native pitch is markedly high in the case non-native speakers in all the three groups.
- The results also indicates that the speakers of all the three language groups are trying to accommodate his/her pitch to suit to a different social group.
- Prosodic differences in the native and English speech by South Indian speakers were detected without annotations of prosodic events.

Imitation for the purpose of integration into a social community by adjusting the speech intonation attributes of fundamental frequency in the non-native English speech has been demonstrated.

7. Conclusions

Following are the conclusions which can be made from comparison between the prosodic characteristics with respect to the pitch frequency contours of non-native speakers corresponding to the three linguistic groups and are found to be different. The results revealed that the F_0 contour in the regional variant of English produced by native Kannada speakers is significantly different from those of other two groups, in terms of the mean and dynamics. Automatic evaluation of non-native speech which includes characteristics of the mother tongue will have better performance over similar algorithms that depend upon target languages. This is particularly true when the text uttered is unknown, present approach allows us to examine the local prosodic changes without incorporating any phonology of the specific languages.

Future investigations are planned for different speech styles, and the composition and characteristics of speaker population to be carefully scrutinised, including multi disciplinary information. Need for inclusions of more number of languages particularly languages of Eastern India. Use of pitch contours to train models of non-native English speech by Indians await future investigations. Also, prosody linked to paralinguistic and non linguistic information need to be analysed.

8. References

[1] L. Neumeyer, H. Franco, V. Digalakis, and M. Weintraub, "Automatic Scoring of Pronunciation Quality", in *Speech Communication*, 2000, vol. 30, pp. 83-93.

[2] F. William, A. Sangwan , and J. H. L. Hansen, "Using Human Perception for Automatic Accent Assessment", in *Interspeech*, pp. 2509-2512, May 2011.

[3] J. Cheng, N. Bojja, and X. Chen , "Automatic Accent Quantification of Indian Speakers of English,", in *Interspeech*, 2013, pp. 2574-2578.

[4] L. Mary and B. Yegnanarayana, "Extraction and representation of prosodic features for language and speaker recognition", *Speech Communication*, vol.50, pp. 782-796, 2008.

[5] G. A. Levow, "Investigating Pitch Accent Recognition in Non-native Speech", *Proceedings of the ACL-IJCNLP conference*, pp. 269-272, 2009.

[6] S. Nisioi, "Feature Analysis for Native Language Identification", in *16th International conference on Intelligent Text processing and computational linguistics*, 2015, pp. 1-15.

[7] A. Ikeno and J. H. L. Hansen, "Perceptual recognition cues in Native English Accent Variation: " Listener Accent and Perceived Accent, and Comprehension" ", in *in IEEE ICASSP,* , 2006, pp. 401-404 .

[8] C. Graham and B. Post, "Second Language acquisition of Intonation: peak alignment in American English", in *Journal of phonetics 66*, 2018, pp. 1-14.

[9] A.Maier, F.Honig, V.Zeissler, A.Batliner, E.Korner, N.Yamanaka, P.Ackermann and E.Noth, "A Language Independent Feature set for the Automatic Evaluation of Prosody", in *Interspeech*, 2009, pp. 600-603.

[10] Ulrike Gut , "Foreign Accent," in *Speaker Classification I. Springer Verlag: Berlin/ Heidelberg, Germany*, 2007, pp. 75-87.

[11] M. Komatsu, "Reviewing human language Identification", *Speaker Classification II,Springer Verlag: Berlin/ Heidelberg, Germany,*, pp. 206-228, 2007.

[12] C. Tiexiera, H. Franco, E. Shriberg, K. Precoda and K. Sonmez, "Prosodic Features for Automatic Text-Independent Evaluation of Degree of Nativeness for Language Learners", in *6th Internation Conference on Speech and Language Processing, 2000*, 4 pages.

[13] X. Wang, K. Evanini, and S. Y. Yoon, "Word Level F0 modeling in Automated Assessment of Non-native read speech", in *Word Level F0 modeling in Automated Assessment of Non-native read speech SLaTE (2015), ISCA workshop on Speech and Language Technology in Education*, pp. 23-27, 2015.

[14] E. Shriberg, L. Ferrer, S. Kajarekar, A. Venkataraman, and A. Stolcke, "Modeling prosodic feature sequences for speaker recognition", in *Speech Communication, 2005*, vol.46, no.3-4, pp. 455-472.

[15] Y. K. Muthusamy, E. Bernard, and R. A. Cole, "Reviewing Automatic Language Identification", *IEEE Signal Processing Magazine,*, vol. 11, No. 4 pp. 33-41, 1994.

[16] C. Lai, K. Evanini and K. Zechner, "Comparative Analysis of Prosodic Features of Native and Non-native Spontaneous Speech", in *New Tools and Methods for Very-Large-Scale Phonetics,* , 2011,

[17] J. P. Arias, N. B. Yoma, H. Vivanko, , "Automatic intonation assessment for computer aided language learning", in *Science Direct, Speech Communication*, 2010, vol. 52, pp. 254-267.

[18] F. Honig, A. Batliner and E. Noth, "Automatic Assessment of Non-native Prosody - Annotation,", in *ISADEPT, 2*, pp. 21-30.

[19] I. Luengo, E. Navas, I. Sainz, I. Saratxaga, J. Sanchez, I. Odriozola and I. Hernaez, "Text Independent Speaker Identification in Multilingual Environments", in *Language Resources and Evaluation Conference (LREC) 2008*, pp. 1814-1817.

[20] K. L. X. Wu and H. Meng, "Intonation Classification for L_2 English speech using multi distribution deep neural net-works", in *Science Direct, Computer Speech and Language 43*, 2017, pp. 18–33.

[21] K. S. Rao and B. Yegnanarayana , "Intonation model-ing for Indian languages", in *Science Direct , Coputer Speech and Language*, 2009, pp. 240-256.

[22] M. P. Harper and M. Maxwell, "Spoken Language Characterization", in *Springer Handbook on Speech Processing and Speech Communication*, 2006, pp. 1-15.

[23] P. Day and A. K. Nandi, "Robust Text-Independent Speaker Verrification Using Genetic Ptogramming", in *IEEE Transactions on Audio Speech, and Language Processing*, 2007, vol. 15(1), pp. 285-295.

[24] G. R. Krishna and R. Krishnan, "Native Language Identification Based on English Accent", in *International Conference on Natural Language Processing (ICON), 2014*, pp. 69-74.

[25] D. Hirst , "Melody Metrics for Prosodic typology: comparing English, French, and Chinese ", in *Interspeech*, 2013, pp. 572-576.

[26] Iskra, D., Grosskopf, B., Marasek, K., Heuvel, H. V. D., Diehl, F., and Kiessling, A., "SPEECON - Speech Databases for Consumer Devices: Database Specification and Validation , in *Proceedings LREC 2002. Third International Conference on Language Resources and Evaluation*, 2002, pp. 329-333.

[27] L. Arslan, J.H.L.Hansen, "Language Accent Classification in American English", in *Speech Communications*, July 1996, vol. 18(4), pp. 353-367.

[28] Doddington G. , "Speaker recognition based on idiolectical differences between speakers," in *In: Proc. EUROSPEECH, Aalborg, Denmark*, 2001, pp. 2521-2524.

[29] M. Kockmann, L. Burget and J. H. Cermocky, "Investigations into prosodic syllable contour features for speaker recogntion", in , *in Acoustics, Speech, and Signal Processing ICASSP,*, 2010, pp. 4418-4421.

[30] J. Cheng, "Automnatic assessment of Prosody in High-Stakes Englsh Test", in *Interspeech, 2011*, pp. 1589-1592.

[31] Eady, S. J., Cooper, W. E. , "Speech intonation and focus location in matched statements and questions", in *Journal of the Acoustical Society of America* , 1980, pp. 402-416.

[32] Fox, R. A, Jacewicz, E., and Hart, J. "Pitch pattern variations in three regional varieties of American English", in *Interspeech, Lyon, France*, 2013, pp. 123-127.

[33] N. Moustroufas, and V. Digalakis, "Automatic pronunciation evaluation of foreign speakers using unknown text," in *Computer Speech and Language*, 2007, vol. 21, pp. 219-230.

[34] M. Farrus, M. Wagner, D. Erro and J. Hernando, "Automatic speaker recognition as a measurement of voice Imitation and conversion ", in *The Intl Jornal of speech, language and the law*, 2010, vol.17.1, pp. 119-142.

[35] D.Crystal, "A Dictionary of Linguistics and Phonetics", in *2nd edition, Oxford: Basil Blackwell, 1985.*

[36] C. Y. Lin and H. C. Wang , "Language Identification using Pitch Contour Information", in *ICASSP 2005*, pp. 601-603.

[37] Wave Surfer , *http://www.speech.kth.se/wavesurfer/*

[38] J. Tepperman and S. Narayanan, "Better Nonna-tive Intonation scores through Prosodic theory," in *INTERSPEECH 2008*, pp. 1813-1816.

[39] N. G. Ward and P. Gallardo, "Non-Native Differeces in Prosodic-Construction Use," in *Dialogue and Discourse*, vol. 8(1), pp. 1-30.