International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 2 296 - 298

Multilingual Speaker Identification using analysis of Pitch and Formant frequencies

Vinay Kumar Jain
Research Scholar
SSTC(SSGI),
Bhilai,India
vinayrich_17@yahoo.co.in

Dr. (Mrs.) Neeta Tripathi
Principal
SSTC(SSEC)
Bhilai,India
neeta31dec@rediffmail.com

Abstract— In the modern digital automated world, speaker identification system plays a very important role in the field of fast growing internet based communications. In India there are many people who are bi-lingual or multilingual, so the requirements to design such system which is used to identify the multilingual speakers. Present paper explores the idea to identify multi-lingual person by basic features. For this the speech signals of three indian languages i.e Hindi, Marathi and Rajasthani are recorded and basic features pitch, first three formant frequency calculated from PRAAT software. The observation has been presented that the pitch and first three formant frequencies F1,F2 and F3 of speaker are increases when speaker change the language from rajasthani to hindi to marathi. The percentage deviation in pitch as well as formant frequencies for Rajasthani and Marathi from hindi are positive and negative respectively for utterance "p". Similar analysis has been perform for 'k aand >. This observation will help to make such system which is used to identify the speaker in multilingual environments.

Keywords- *Pitch, formant frequencies, Multilingual Speaker, PRAAT, Percentage deviation.*

*****

I. INTRODUCTION

Today's era the speakers knows more than one languages and it is required that the speaker identification system should give the better performance for this type of speakers who are speaking multiple languages. When the speaker recognition is being transferred to real applications, the need for greater adaptation in recognition is required. The performance of the monolingual speaker identification systems tends to decreases when speaker is speaking in another language. Therefore we need to make such systems which can work for multiple languages.

Languages are usually influenced by other languages that are present in the environment and by the speaker's mother tongue[2]. Multilingual speech processing (MLSP) is a distinct field of research in speech and language technology that combines many of the techniques developed for monolingual systems with new approaches that address specific challenges of the multilingual domain [8].

In order to find some statistically relevant information from speech signal, it is important to have mechanisms for reducing the information of each segment in the audio signal into a relatively small number of parameters, or features. Feature extraction is the first step for the multilingual speaker identification system. Many algorithms were suggested/developed by the researchers for feature extraction. The basic features i.e. Pitch and first three formant frequencies will help for identification of person on multilingual base

II. DATABASE GENERATION

For multilingual speaker identification system, the database of different speakers has been recorded.The sampling rate of recorded sentences is 44KHz.The segmentation is done by Goldwave software. The sentences consist consonants i.e च ष , झ has been considered for the recording. The total number of speakers involved are 20 including males and females. The recorded sentences are:

मुझे चाय पीना पसंद है। चाय मे षक्कर कम है। तिरंगा हमारा झंडा है।
मला चाय पसंद आहै चाय मधे षक्कर कमी आहै तिरंगा अमच्चये झंडा आहै
मन्ने चाय पीनी पसंद है। चाय मो षक्कर कम है। तिरंगा मारा झंडा है।

III. FEATURE EXTRACTION

The original speech signal contains redundant information. For speech recognition and application eliminating such redundancies helps in reducing the computational overhead and also improve system accuracy. Therefore all most speech application involves the transformation of signal to set of compact speech parameter.

In this work mainly pitch and first three formant frequencies F1,F2 and F3 are calculated from the speech signal of different languages. The pitch is fundamental frequency F0 and it is determined by the vibratory frequency of the vocal folds. The standard range of pitch is from 75 Hz to 500 Hz for human voice. Formants are frequency peaks which have, in the spectrum, a high degree of energy. They are especially prominent in vowels. Each formant corresponds to a resonance in the vocal tract and the spectrum has a formant at approximate every 1000 Hz.

IV. RESULT AND DISCUSSION

In this paper, investigation has been made for two basic features pitch and first three formant frequencies F1,F2 and F3 for male and female speakers in three languages Hindi and Marathi and Rajasthani. The analysis is done for the three utterance च , ष and झ in three languages Hindi and Marathi and Rajasthani. Base of the analysis is to observe the variation in pitch and first three formant frequencies F1,F2 and F3 if the

 ISSN: 2321-8169
296 - 298

speaker change the spoken language. The following observations were recorded in table-1 and table-2.

Table-1: Pitch and percentage deviation for utterance च.

Speakers	Pitch(Hz)			% Deviation of Hindi to Marathi	% Deviation of Hindi to Rajasthani
	Hindi Language	**Marathi Language**	**Rajasthani Language**		
1	**160**	**168**	**142**	-5.00	11.25
2	**155**	**189**	**149**	-21.94	3.87
3	**139**	**148**	**125**	-6.47	10.07
4	**169**	**187**	**151**	-10.65	10.65
5	**142**	**147**	**140**	-3.52	1.41
6	**122**	**165**	**120**	-35.25	1.64
7	**141**	**163**	**125**	-15.60	11.35
8	**124**	**155**	**121**	-25.00	2.42
9	**133**	**173**	**125**	-30.08	6.02
10	**144**	**165**	**132**	-14.58	8.33
11	**126**	**155**	**125**	-23.02	0.79
12	**137**	**169**	**125**	-23.36	8.76
13	**160**	**170**	**135**	-6.25	15.63
14	**168**	**173**	**145**	-2.98	13.69
15	**144**	**150**	**129**	-4.17	10.42
16	**143**	**186**	**122**	-30.07	14.69
17	**146**	**157**	**136**	-7.53	6.85
18	**125**	**172**	**120**	-37.60	4.00
19	**145**	**163**	**139**	-12.41	4.14
20	**132**	**155**	**122**	-17.42	7.58

From Table-, it has been observed that when speaker change the language, the pitch value has changed for च. The pitch values of Marathi language are more as compared to hindi and Rajasthani languages. The percentage deviation Marathi language and rajasthani languages from hindi language is presented in table no-1.

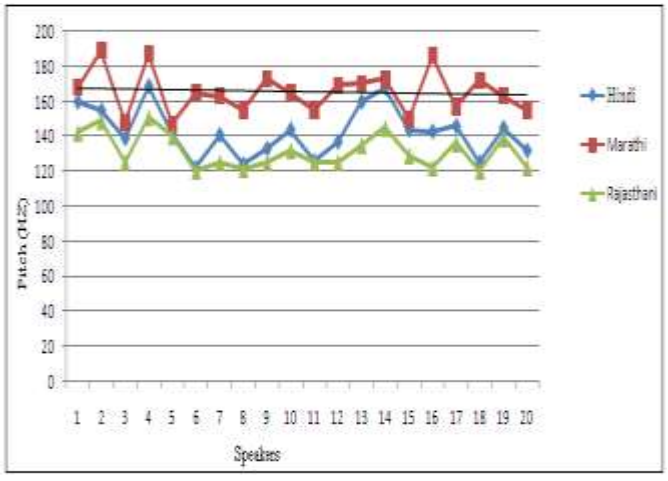

Figure-1:Variation of pitch for three languages of च

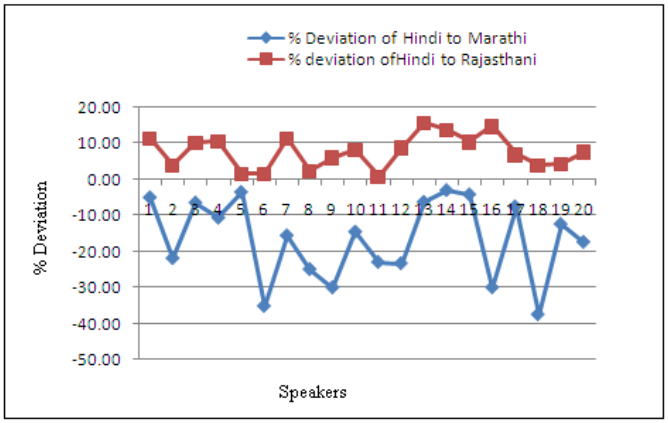

Figure-2: Percentage deviation of Marathi and Rajasthani Languages from Hindi of च

The variation of pitch of च in three languages as shown in figure-1 The observation has been presented that the pitch of speaker are increases when speaker change the language from rajasthani to hindi to marathi. Therefore the percentage deviation in pitch for Rajasthani and Marathi from hindi are positive and negative respectively for utterance "च" which is shown in figure-2. Similar analysis has been perform for ष and झ .

Table-2: % deviation of formant frequencies (F1,F2 & F3) for utterance च

Speaker	F1		F2		F3	
	% Deviation (Hindi to Marathi)	% Deviation (Hindi to Rajasthani)	% Deviation (Hindi to Marathi)	% Deviation (Hindi to Rajasthani)	% Deviation (Hindi to Marathi)	% Deviation (Hindi to Rajasthani)
1	-8.30	18.49	-5.45	17.16	-2.45	6.13
2	-10.19	14.79	-8.69	3.47	-5.12	5.00
3	-0.64	37.09	-4.76	4.85	-2.48	5.51
4	-11.61	28.08	-5.98	1.93	-3.51	4.10
5	-13.70	5.26	-3.27	5.93	-2.77	3.76
6	-7.46	22.71	-0.50	16.89	-7.81	6.00
7	-8.35	17.85	-2.88	11.57	-2.68	5.68
8	-19.35	10.88	-1.92	6.71	-4.56	16.96
9	-19.11	2.78	-4.85	7.25	-8.78	2.71
10	-11.63	14.51	-10.08	16.48	-7.10	4.92
11	-13.07	14.47	-3.36	10.76	-2.01	6.70
12	-11.40	27.43	-6.66	7.91	-5.03	17.07
13	-55.56	11.93	-9.31	3.41	-4.75	4.75
14	-36.29	4.21	-8.95	8.53	-7.16	4.13
15	-10.93	21.34	-2.77	2.66	-4.24	5.08
16	-24.68	12.03	-2.06	0.36	-2.40	18.29
17	-10.96	16.67	-0.56	3.06	-3.45	8.41
18	-10.55	1.23	-6.26	11.76	-2.13	15.99
19	-12.38	17.65	-9.73	1.33	-4.37	5.97
20	-22.42	5.95	-9.94	11.01	-9.41	6.88

In Table-2 the percentage deviation of first three formant frequencies F1,F2 and F3 has been presented. The percentage deviation shows that when the speaker change the language , the first three formant frequencies F1,F2 and F3 has changed. It has been observed that the percentage deviation in formant frequencies F1,F2 and F3 for Rajasthani and Marathi from hindi are positive and negative respectively for utterance "च". Similar analysis has been perform for ष and झ.

V. CONCLUSIONS

Feature extraction has the ability to improve the performance of multilingual speaker identification system. Presented observation, the pitch and first three formant frequencies F1,F2 and F3 of the speakers have changed when the speaker change the spoken language. The pitch and first three formant frequencies F1,F2 & F3 of Marathi lingual speakers has more as compared to hindi and rajasthani lingual speakers. This observation will help to make such system which is used to identify the speaker in multilingual environments.

REFERENCES

[1] S. Agrawal, and et al,. "Prosodic feature based text dependent speaker recognition using Machine learning althorithms". International Journal of Engineering Science and Technology.2(10), 2010:5150-5157.

[2] V Alabau. and Mart´ınez C. D. "Bilingual Speech Recognition in two phonetically similar Languages". Jornadas en Tecnologia del Habla, Zaragoza.IV, 2006,197-202.

[3] W.Bharti. and et al. "Marathi Isolated Word Recognition System using MFCC and DTW Features".ACEEE Int. J. on Information Technology, Vol. 01, No. 01, Mar 2011:21-24

[4] U.Bhattacharjee. and K. Sarmah, "A multilingual speech database for speaker recognition", IEEE International Conference on Signal Processing, Computing and Control (ISPCC), Waknaghat Solan,2012:1-5.

[5] U.Bhattacharjee. and K. Sarmah. "Development of a Speech Corpus for Speaker Verification Research in Multilingual Environment". International Journal of Soft Computing and Engineering (IJSCE).2(6):2012,443-446.

[6] U.Bhattacharjee. and K. Sarmah. "GMM-UBM Based Speaker Verification in Multilingual Environments". IJCSI International Journal of Computer Science Issues.9(6), 2012:373-380.

[7] R Ranjan, and et al,. "Text-Dependent Multilingual Speaker Identification for Indian Languages Using Artificial Neural Network".3rd International Conference on Emerging Trends in Engineering and Technology.Goa, India,2010:632-635.

[8] H. Bourlard.,Dines and et al, **"**Current trends in multilingual speech processing". Academy of Sciences Sadhana . 36(5),2011:885–915.

[9] M. A. Carl, "Multilingual and Crosslingual Acoustic Modelling for Automatic Speech Recognition". Ph.D. Thesis. University of South Florida, 2008.

[10] K.Umapathy, and et al. "Audio Signal Feature Extraction and Classification Using Local Discriminant Base",. IEEE Transaction on Audio, Speech, and Language Processing.15(4), 2007:1236-1246.

[11] Desmet T. and Duyck W. "Bilingual Language Processing", Language and Linguistics Compass 1/3, 2007:168–194.

[12] M Shaneh, and A Taheri. "Voice Command Recognition System Based on MFCC and VQ Algorithms". World Academy of Science, Engineering and Technology-33, 2009,534-538

[13] F. Diehl, "Multilingual and Crosslingual Acoustic Modelling for Automatic Speech Recognition". Ph.D. Thesis,Departament de Teoria del Senyal i Comunicacions Universitat Polit`ecnica de Catalunya, 2007.

[14] G.Kaur and H.Kaur, "Multi Lingual Speaker Identification on Foreign Languages Using Artificial Neural Network with Clustering". International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 5, May 2013 ISSN: 2277 128X:14-20.

[15] M.Ferras, and et al,. "Comparison of Speaker Adaptation Methods as Feature Extraction for SVM-Based Speaker Recognition". IEEE Transaction on Audio, Speech, and Language Processing.18(6), 2010:366-1378.

[16] H. A.Patil and et al "Design of Cross-lingual and Multilingual Corpora for Speaker Recognition Research and Evaluation in Indian Languages". International Symposium on Chinese Spoken Languages Processing(ISCSLP 2006),Kent Ridge, Singapore.

[17] V. Tiwari, "MFCC and its applications in speaker recognition". International Journal on Emerging Technologies 1(1), 2010: 19-22.

[18] D. Imseng. "Multilingual speech recognition A posterior based approach". Ph.D. Thesis, Ecole Polytechnique Federale De Lausanne,2013

[19] D. Lyu and et al, "Acoustic Model Optimization for Multilingual Speech Recognition". Computational Linguistics and Chinese Language Processing.13(3), 2008: 363-386..

[20] L.R Rabiner and R.W Schafer. Digital Processing of Speech Signals: Pearson Education, 9th edition. ISBN No.: 978-813-317-0513-1.