

Accent Classification among Punjabi, Urdu, Pashto, Saraiki and Sindhi Accents of Urdu Language

Afsheen, Saad Irtza, Mahwish Farooq, Sarmad Hussain
 firstname.lastname@kics.edu.pk, afsheen@kics.edu.pk, saad.irtaza@kics.edu.pk
 Center for Language Engineering, Al-Khawarizmi Institute of Computer Science
 University of Engineering and Technology, Lahore, Pakistan

Abstract

Automatic Speech Recognition (ASR) is a key component in Human Computer Interaction (HCI) applications. Stability of ASR systems largely depends on accent, gender and age of speakers. In this paper, a study has been conducted to classify five different accents of Urdu language spoken in Pakistan i.e. Punjabi, Urdu, Pashto, Saraiki and Sindhi. The five accents have been classified using feature formants (F1 and F2) and Mel frequency cepstral coefficients (MFCCs).

Introduction

The objective of an automatic speech recognition system is to convert speech into text. The performance of ASR systems depends on the training data on which acoustic models have been trained. Speaker dependent (SD) ASR systems perform better than speaker independent (SI) ASR systems. The acoustic variations in speakers are mainly due to age, gender and regional accents. Among these factors, accent is the most leading factor that contributes to a higher error rate in ASR systems. Accent refers to the articulation pattern that a speaker follows to produce a particular sound and it is also related to speaker's first language, which affects speaker's perception and production of speech. Dialect and accent do not cause much trouble when communicating among humans but gives poor recognition results with ASR systems. In different geographical regions of Pakistan, 59 languages are spoken. Based on the native language, there are six prominent accents used in Pakistan, namely Urdu, Punjabi, Pashto, Saraiki, Sindhi and Balochi. Although less than 8% of Pakistanis speak Urdu as their first language, but it is spoken and understood as a second language by almost all Pakistanis. In learning a second language, vowel articulation is the key factor to take care of. Therefore, in our experiments we use vowels to classify five different accents of Urdu language.

Methodology

Two experiments have been conducted to classify Punjabi, Urdu, Pashto, Saraiki and Sindhi accents of Urdu language based on following acoustic features.

- 1- Formant frequencies (F1 and F2)
- 2- Mel frequency cepstral coefficients (MFCCs)

In experiment 1, formants are calculated from the midpoint of each vowel and their mean and standard variation values are used for classification while in the second experiment, Euclidian distance is used to calculate acoustic distance between different accents of Urdu language. The Euclidian distance formula used to calculate acoustic distance between Punjabi and Urdu accents of Urdu language based on formants is as follows:

$$d_{(Pun,Urd)} = \sqrt{(F_1^{PUN} - F_1^{URD})^2 + (F_2^{PUN} - F_2^{URD})^2}$$

The Euclidian distance formula used to calculate acoustic distance between Punjabi and Urdu accents of Urdu language based on MFCCs is as follows:

$$d_{(Pun,Urd)} = \sqrt{(mfcc_1^{PUN} - mfcc_1^{URD})^2 + \dots + (mfcc_7^{PUN} - mfcc_7^{URD})^2 + \dots + (mfcc_{12}^{PUN} - mfcc_{12}^{URD})^2}$$

The training and testing corpus consists of 760 utterances from native speakers of above mentioned five languages. The corpus used for two experiments is as follows:

Vowel	IPA	Punjabi	Urdu	Pashto	Saraiki	Sindhi
A	ə	276	109	295	80	142
AA	ɑ:	255	82	164	51	90
AE	e:	69	21	37	20	20
I	i	125	64	52	12	29
II	i:	65	25	57	19	30
OO	o:	39	16	15	7	15
U	u	62	28	39	14	20
UU	u:	36	23	22	8	13

Table 1: Vowel data utterance

Classification Results

Experiment 1:

The following graphs show the classification results of Experiment 1 for above listed eight vowels of Punjabi, Urdu, Pashto, Saraiki and Sindhi accents.

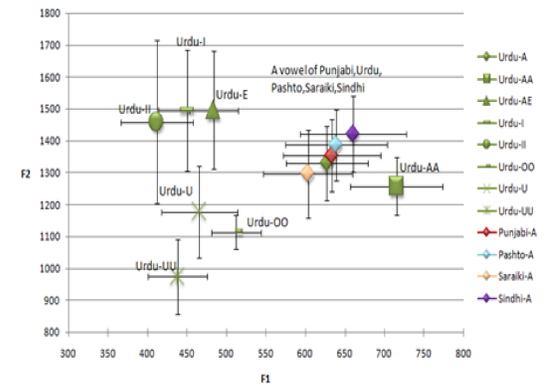


Figure 1: Comparison of A vowel of 5 accents and all vowels of Urdu

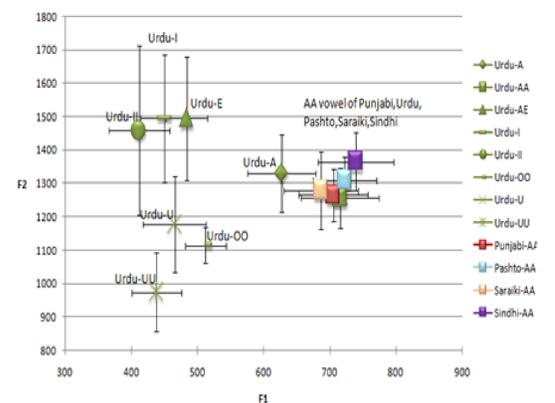


Figure 2: Comparison of AA vowel of 5 accents and all vowels of Urdu

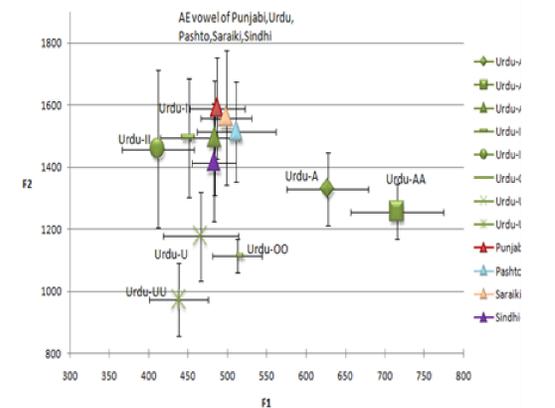


Figure 3: Comparison of AE vowel of 5 accents and all vowels of Urdu

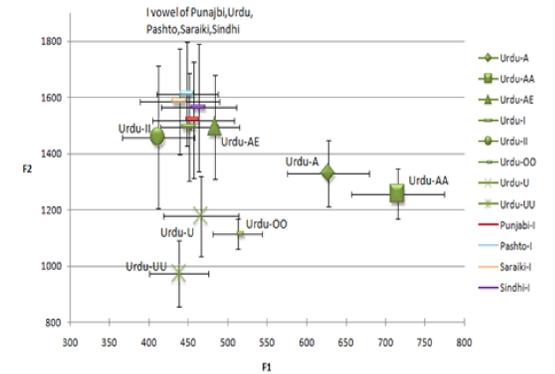


Figure 4: Comparison of I vowel of 5 accents and all vowels of Urdu

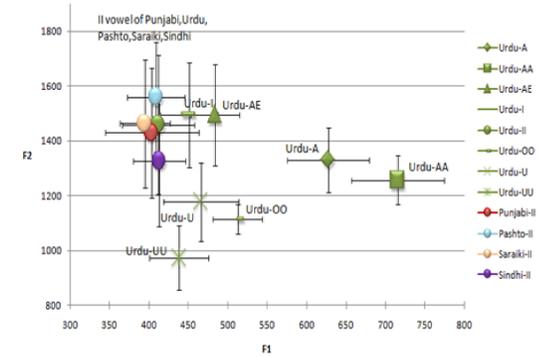


Figure 5: Comparison of II vowel of 5 accents and all vowels of Urdu

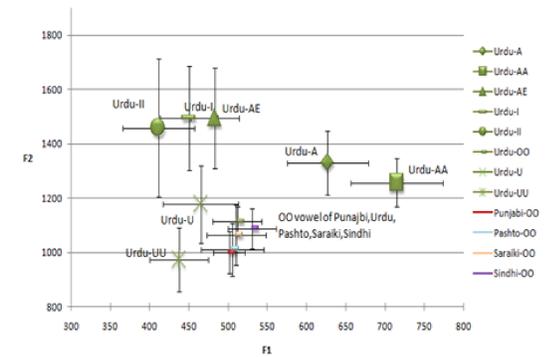


Figure 6: Comparison of OO vowel of 5 accents and all vowels of Urdu

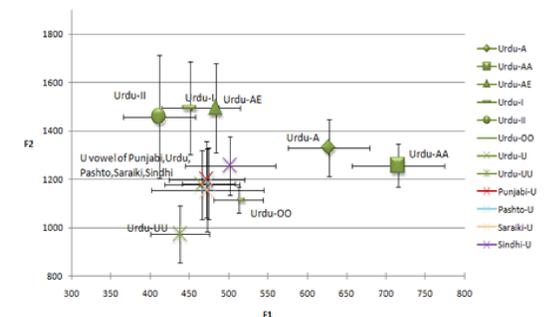


Figure 7: Comparison of U vowel of 5 accents and all vowels of Urdu

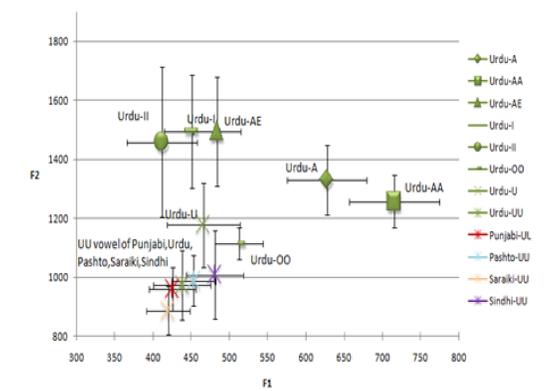


Figure 8: Comparison of UU vowel of 5 accents and all vowels of Urdu

Experiment 2:

In Experiment 2, distances calculated by Euclidian formula have been used to classify accents of Urdu language. Table 2 shows the distance between different accents, computed by summing respective distances of these accents over all vowels.

Accents	MFCC	Formants
Punjabi-Urdu	0.637308	0.1137
Punjabi-Pashto	0.713092	0.287106
Punjabi-Saraiki	1.47057	0.426182
Punjabi-Sindhi	1.394903	0.392127
Urdu-Pashto	0.648211	0.38703
Urdu-Saraiki	1.514395	0.402592
Urdu-Sindhi	1.324604	0.499803
Pashto-Saraiki	1.613643	0.39677
Pashto-Sindhi	1.317765	0.300893
Saraiki-Sindhi	2.263357	0.670466

Table 2: Distance between different accents of Urdu language based on Formants and MFCCs

Discussion

The figures 1 to 8 of experiment 1 show that using F1 and F2, all vowels of the Urdu accent can be classified, with an exception of short vowel I which is being confused with long vowels II and AE of Urdu language (shown in all figures 1 to 8). But on the other hand it is very clear from the above graphs that using F1 and F2, it is difficult to classify Punjabi, Urdu, Pashto, Saraiki and Sindhi accents of A, AA, AE, I, II, OO, U and UU vowels of Urdu language.

Table 2 of experiment 2 shows that distance calculated between two accents using MFCCs is always greater than the distance calculated using formant frequencies. Therefore, it can be concluded that on the basis of distance calculated using MFCCs, the probability of an accent to get confused with other accents is minimal. As given in table 2, the calculated distance between "Urdu and Saraiki" accents based on formant frequencies values is 0.402592 while based on MFCC values is 1.514395 which is almost three times greater than that calculated using formant frequencies. Therefore, MFCCs can be used to classify Punjabi, Urdu, Pashto, Saraiki and Sindhi accents of Urdu language.

Conclusion

Above mentioned results of the two experiments show that two dimensional formant features, F1 and F2, are not sufficient to classify Punjabi, Urdu, Pashto, Saraiki and Sindhi accents of Urdu language spoken in different geographical regions of Pakistan. Therefore, there is a need to explore more dimensions of speech data. This need has been accomplished by using twelve dimensional Mel-frequency cepstral coefficients (MFCCs), that retain accent related information of the speaker, to classify above mentioned five accents. The results of the two experiments show that MFCC vectors can be used to classify Punjabi, Urdu, Pashto, Saraiki and Sindhi accents of Urdu language.