An Acoustic Phonetic Study of Six Accents of Urdu in Pakistan

by

Mahwish Farooq

M. Phil in Applied Linguistics

Department of English Language and Literature

School of Social Sciences and Humanities

University of Management and Technology

2014

An Acoustic Phonetic Study of Six Accents of Urdu in Pakistan

by

Mahwish Farooq

M. Phil in Applied Linguistics

This research was submitted to University of Management and Technology, Johar Town, Lahore in the partial fulfillment for the requirements of the Degree

of

M. Phil in Applied Linguistics

Department of English Language and Literature

School of Social Sciences and Humanities

University of Management and Technology

2014

ACKNOWLEDGEMENTS

Start with the name of Allah Almighty who holds and guides us in darkness. I owe to start with the praise of Allah Almighty as I know I am nothing without His help.

It is my pleasure to acknowledge different people who have been contributed a lot to my thesis. Firstly, I am incredibly thankful to my supervisor Prof. Dr. Sarmad Hussain for inspiring me with the idea for this project and of course for supervising me. I am very lucky to have a supervisor like him. I would like to express my gratitude for his patience, encouragement and guidance during my research. He is an inspiration for me to enter into the challenging field of research. I am very obliged and thankful to him for responding my queries each and every time during this work.

I am very grateful to the former chairperson Mr. Rao Jalil and the present chairperson Dr. Shaban of Department of English Language and Literature, for their co-operation and guidance. I am thankful to all my teachers who helped me in my queries. Specially, I would like to thank Prof. Dr. Abdul Hameed, dean of School of Social Sciences and Humanities who has guided me in analyzing the data statistically well. I am also grateful to Ms. Sana and Sr. Asad who always encouraged and motivated me. I would like to say special thanks to my friends Afsheen Rafaqat Ali and Saba Noman for guiding me in the technical problems. I am grateful to all my friends and colleagues who helped me in my queries.

Finally, I am indebted to my beloved parents and my family from my core of the heart. Without their affectionate love, continuous encouragement and support, it would be very difficult for me to complete my degree and especially this research work.

ABSTRACT

Urdu is a lingua franca, an official language (Mahmood, 2004) and the mother tongue of only 7.5% population in Pakistan (Zia, 2011). The present study deals with an acoustic phonetic analysis which is conducted for finding out the accent variation in Pakistani Urdu. As, we know, Pakistan is a multilingual country therefore the purpose of the research is to analyze the influence of the other languages on Urdu. This research is based on quantitative methodology. The list of the 139 district names has been used as a corpus. The recording of the utterances by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages have been collected by developing a system for online data collection. 30 Volunteers of each language have been enrolled to provide recordings for speech corpus. The corner vowels from the utterances have been evaluated by comparing them with the acoustic properties. Accent differences are acoustic manifestations of differences in duration, pitch, intonation pattern and of course the difference in phonetic transcription (Yan & Vaseghi, 2002). As this is the first step in the work of an acoustic phonetic study of Urdu accents therefore only variations of vowels have been identified by measuring the formant frequencies manually in PRAAT software. The preliminary analysis of first and second formant frequencies showed the differences in the characteristics of vowels. This research has verified that the formant frequencies of the vowels (uttered by the speakers of six major languages) show differences and variations across phonetic context. This is due to the fact that Urdu is the second or third language for Pakistani speakers (Rehman, 2002). It has also verified that some utterances are showing more similar values of formant frequencies (of the speakers' utterances) than the others e.g. the formant frequencies in the utterances of Urdu and Punjabi speakers while

the other showing more variant and dissimilar formant tendencies as in the case of Urdu, Punjabi and Sindhi language speakers' utterances.

Table of Content

Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Pakistan is a multilingual country as people could communicate and understand more than two languages but the majorly spoken languages are only six those are Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. Among the different linguists, there is a controversy in deciding the total number of languages in Pakistan. The difference in the number of spoken languages occurs because of the confusion of languages as dialects or independent languages as Rehman (2002) said that there are 59 spoken languages in Pakistan whereas according to Ethnologue, there are 72 spoken languages and most of these languages are belonging to Indo-Aryan, Indo-Iranian, Indo-European, and Turkic language families (Gordon, 2005).

Even though, Urdu is an official language of Pakistan (Mahmood, 2004) but the accent, lexicon and prosody vary across geographical areas. In Pakistan, there are six prominent accents of Urdu, based on the native languages of the speakers; namely Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki (Rehman, 2002). Accent variation is really based on the duration, pitch, amplitude and the intonation patterns of phonemes (Yan et al., 2005; Reetz & Jongman, 2008). The term “accent” means a pattern of pronunciation used by a speaker to which he or she belongs. Linguistically, accent variation does not only lie in phonetic characteristics but also in prosody (Cole, 2012). Accent is also one of the main factors that impacts Automatic Speech Recognition (ASR) system (Benzeghiba, et al., 2006).

This research is a first step in analyzing the accent variation in Pakistani Urdu language. Therefore, in this study, only vowels have been analyzed acoustically by measuring formant frequencies; F1, F2. The formant frequencies greatly contribute to the better understanding of spoken languages and their accents. Formant frequencies of vowels have been analyzed in PRAAT and mean values have been calculated for each vowel. Then, these mean values have been compared pair wise and group wise. This analysis has verified that multilingualism is the major cause of accent variation in Urdu language and also has proved that Pakistani Urdu has six major accents.

1.1. Scope of the Study

The acoustic phonetic study of vowels is beneficial for Human Computer Interaction (HCI) used for industrial and social applications i.e. applications for telephonic communication, banking, military, health care and education. However, significant development and research have been achieved for English language. Hence, the present study focuses on the progress and development of Urdu language, especially in the area of acoustic phonetics that will investigate the physical properties of sound and also will prove beneficial for the Automatic Speech Recognition (ASR) system.

1.2. Statement of the Problem

A lot of work has been done in Urdu literature and grammar. But there is no major work done on the acoustic properties of the different accents of Urdu. Pakistani speakers use Urdu for mutual communication but they may have different first and second languages. Accents of the different native languages affect the accent of Urdu as

well. This acoustic analysis has been conducted for finding the similarities and differences among major accents of Urdu.

1.3. Motivation and Goal of the Study

In the constitution of 1956, Urdu has been declared as a national language (Rehman D. T., 2006) and is also a lingua franca in Pakistan. For that reason, it is widely used, both formally and informally; in communication, education, literature, business, media and governance.

Urdu is also the first language of only 7.5% population in Pakistan: the reason is, in different districts of Pakistan, people have different mother languages but Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki have comparatively larger population than other languages and all these languages may or may not be intelligible among people in different geographical areas (Zia, 2011) as it is evidently clear that each language has region specific prosodic and lexical features (Gordon, 2005).

But Urdu is the second or third language for almost all Pakistani speakers therefore, it has acquired Pakistani flavor by absorbing Punjabi, Sindhi, Balochi, Pashto and Saraiki languages in terms of accents and vocabulary so there are six major accents of Urdu in Pakistan (Rehman T. , 2002).

The present study will also proved beneficial for the curriculum development of Urdu in Pakistan. And the received phonetic inventory could be added in Urdu grammar and composition books for the development and progress of the spoken Urdu language in Pakistan. The phonetic inventory will also prove beneficial for teachers' training. The

trained teachers will ultimately play their role in the progress and development of Urdu language among young learners.

1.4. Research Question

This study deals with the acoustic phonetic analysis of the vowels. It presents the answer to the question:

Whether all the six major languages of Pakistan affect Urdu language or not?

1.5. Delimitations

This proposed study only deals with the segmental features, not with the supra-segmental features of the corner vowels.

CHAPTER 2

LITERATURE REVIEW

The research is conducted for finding out the accent variations in Urdu, the national language of Pakistan. The aim of this project is the production of *Urdu Phonological Atlas* which will be explained the present condition of the phonological system of accents in Pakistan. Medium of mobile or telephone had been used for the collection of data. All the native speakers who were the members of the local speech community were selected from different universities of Pakistan for the collection of Urdu speech data. The speech data had been analyzed acoustically in PRAAT. This study will be the first research on the acoustic phonetic analysis of Urdu accents in Pakistan. It presents the answer of the question:

1. What are the majorly spoken accents of Urdu language in Pakistan?

This research provides a first overview of the accent variations; affecting Urdu language in Pakistan which is the principal aim of this project. The current state of Urdu is the result of the active process of changes and diversifications due to the affect of other major neighboring languages which we have been tracing. As the focus of the present research is the comparison of the utterances of Urdu speakers with the utterances of Punjabi, Sindhi, Balochi, Pashto and Saraiki language speakers. This study has only dealt with the phonemic information of sounds therefore the orthography of these languages would not be discussed here. As, a spoken language consists of successions of sounds emitted by the speech organs along with certain attributes: each speech sound belongs

either to the vowel or the consonant. It was also discovered that there are different sounds which are neither vowels nor consonants. Thus, may be termed as semi-vowels /w/, /r/, /j/ and laterals /l/ by showing the mixed properties of vowels and consonants (Roach, 2009).

Important Terms

Articulatory phonetics, auditory phonetics, acoustic phonetics, formant features (F1, F2 values), PRAAT, spectrogram, phonetic inventory, accents and multilingualism

2.1. Articulatory Phonetics

Articulatory phonetics deals with the movement of different parts of vocal tract for the speech production (Universitat Bielefeld, 2010). Articulatory phonetics also tells us about the places and manners of articulation for different phonemes. The articulation of speech is based on the following three processes;

a) the airstream process,
b) the phonation process,
c) the oro-nasal process

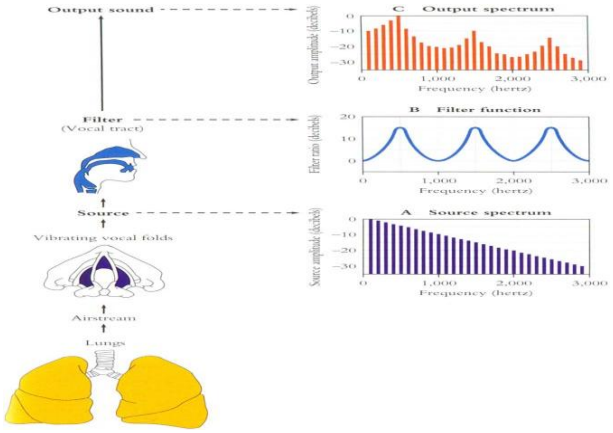


Figure 1: Process of Articulation[1]

[1] Retrieved 12-8-2013: http://www.uni-bielefeld.de/(en)/lili/personen/dschlangen/publications.html

2.2. Auditory Phonetics

Auditory phonetics[2] deals with the perception of sounds or the way in which sounds are being heard and perceived. If articulatory phonetics deals with the speaker than auditory phonetics deals with the listener.

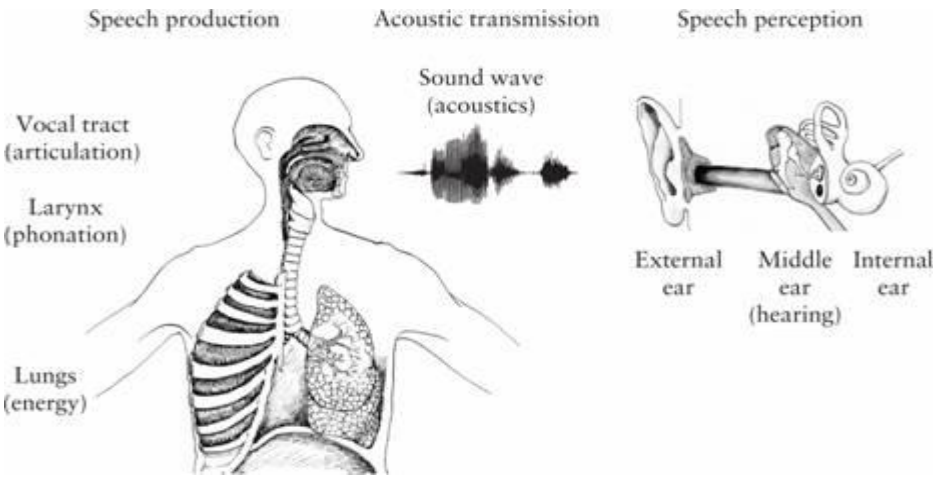


Figure 2: Production and Perception of Speech

It directly deals with the physiology and anatomy but directly related with the utterance of speech. Anatomy and physiology are distinct processes but are closely interrelated with each other. One is the audition or perception of the sounds by auditory apparatus; sending it to brain by converting it into neural signals and the other deals with the analysis of these neural signals in the brain by decoding and understanding of the verbal information given by the speaker. The first anatomical organ is the ***ear*** of the listener and the last is the ***brain*** of the listener so we can say that the speech production is a ***chain process*** which starts from the speaker's brain and ultimately ends to the listener's brain (Universitat Bielefeld, 2010).

[2] Retrieved 12-8-2013: http://www.ebooks.unibuc.ro/filologie/mateescu/pdf/23.pdf

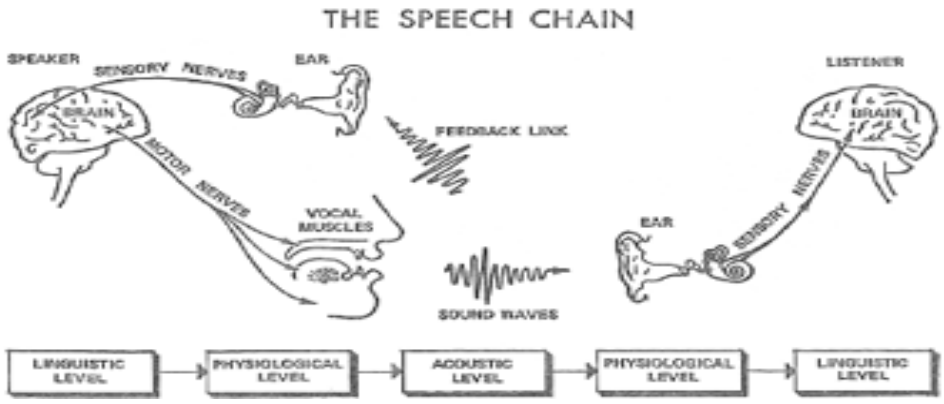


Figure 3: The Speech Chain, speaker's mind to listener's mind[3]

2.3. Acoustic Phonetics

The acoustic phonetics deals with the physical properties (pitch, loudness, amplitude, quality and spectrographic properties) of the sound waves during the speech production process (Universitat Bielefeld, 2010).

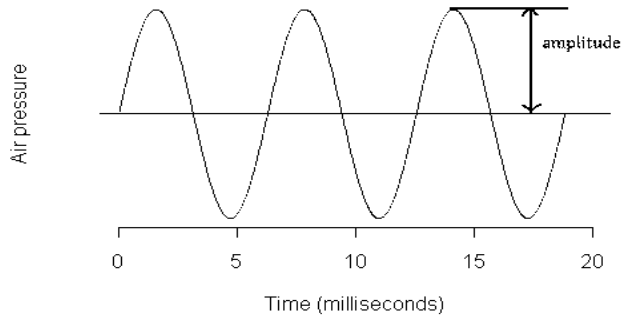


Figure 4: Physical Properties of Sound Wave[4]

2.4. PRAAT

PRAAT is a computerized program used to analyze the sound waves by playing, annotating and analyzing the sound objects in terms of acoustic properties i.e. frequency, pitch, intensity, etc. Formants are shown in red dotted lines in spectrogram of PRAAT.

[3] Retrieved 12-8-2013: http://www.uni-bielefeld.de/(en)/lili/personen/dschlangen/publications.html
[4] Retrieved 12-8-2013: http://www.uni-bielefeld.de/(en)/lili/personen/dschlangen/publications.html

Different tiers are used to segment the speech waveform for further analyses (Universitat Bielefeld, 2010).

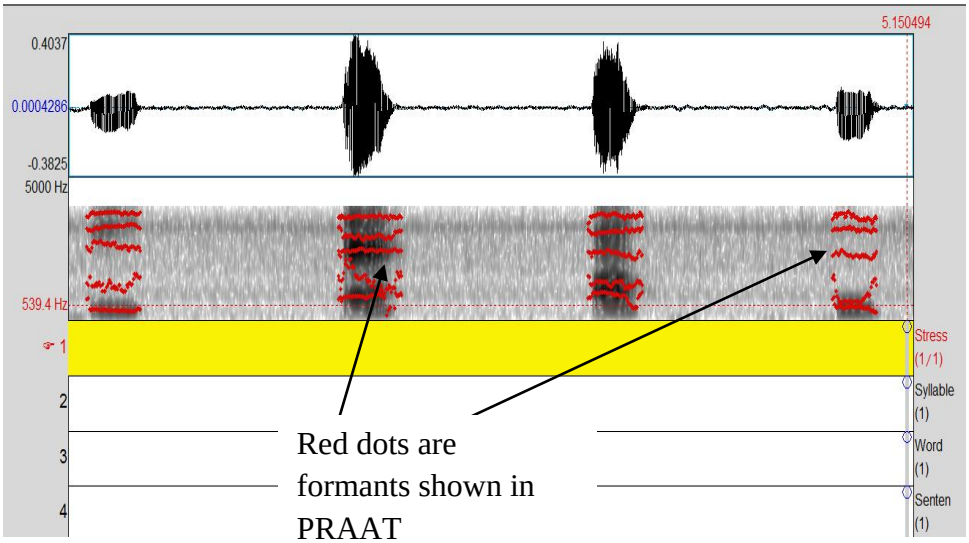


Figure 5: Formants Analysis in PRAAT

2.5. Spectrogram

Spectrogram is the graphical representation of sound waves by explaining their component frequencies. It also shows three dimensional information i.e. frequency (vertical axis), time (horizontal axis) and acoustic energy means the formant frequencies; the dark shading bands on spectrogram (Universitat Bielefeld, 2010).

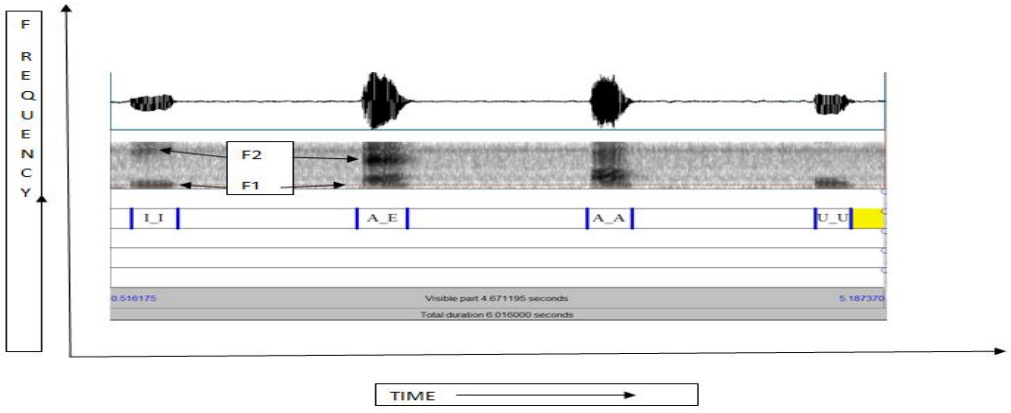


Figure 6: Three Dimensional Spectrograms

2.5.1. Formant Frequencies

Formants are produced with the vibrations of air stream in the vocal tract, and these vibrations may or may not be harmonic; in other words, formants are the concentration of the acoustic energy and groups of overtones corresponding to a resonating frequency of the air in the vocal tract and also are the energy peaks which determine the quality of sound waves especially vowels therefore, vowels are characterized by using three formants; F1, F2 and F3 (Karamat, 2012: Universitat Bielefeld, 2010).

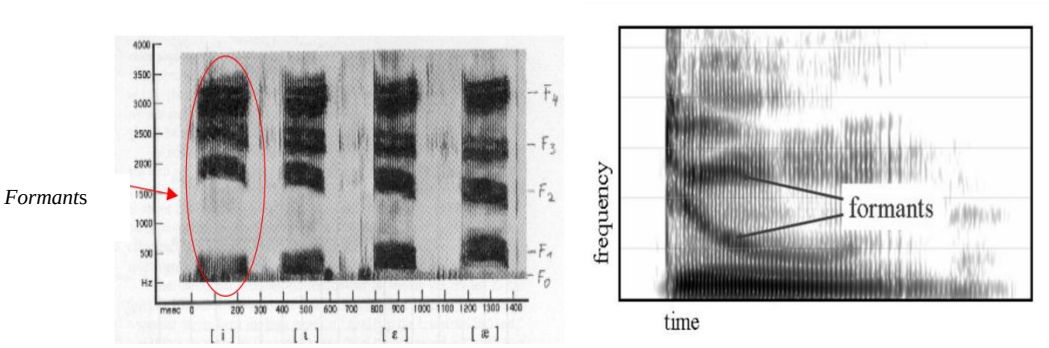


Figure 7: Spectrogram showing Formants

2.6. Phonetic Inventory

According to Lan Maddieson[5], a phonetic inventory could explain all speech sounds, regardless of this, the sounds are produced correctly or not in a specific language or phonetic inventory is a key which explains the different phonemic sounds in a specifically given language where the set of consonants and vowels are called consonantal and vocalic inventories respectively.

[5] Retrieved 12-8-13: http://www.wals-online-chapter-consonant-inventories

There is an acoustic hypothesis for checking the phonetic inventory in the connected speech of a given language. According to this hypothesis, there are many acoustic parameters for the detection of consonants and vowels e.g. amplitude, periodicity and spectrum along with some primary features e.g. voicing, continuant, grave, compact, nasalization, palatalization, pharyngealization, clicks, etc. These auditory features are strong and can be detected in a short period of time. The length is less distinctive feature in the consonants and vowels. Further different less common features among consonants are: discontinuity of air stream, voicing, glottalization, aspiration, voice quality and affrication. This is also applicable to all vowel inventories. The modal vowel inventory is /ɪ, e, a, o, u/ which is present almost in all languages.

2.6.1. Consonants

In 2009, Roach defines different types of consonants in following ways; there are many types of consonants and all have in common is that they obstruct the flow of air in vocal or nasal tract completely or partially. Different consonants behave differently: those which make the maximum obstruction are the *plosives* by making a complete stoppage of the air stream in oral cavity. *Nasal* consonants produce with complete closure of the oral cavity but are less obstructive than plosives and air is released through the nose. *Fricatives* make a considerable restriction to the flow of air but not a complete stoppage. *Laterals* stop the air flow only in the center of the oral cavity, not at the sides and cause minor air closure. Other sounds are known as *approximants;* make very slight obstruction to the flow of air because of which they could almost be thought as semi vowels especially in English.

2.6.2. Vowels

Vowels are "the sounds in which there is no obstruction to the flow of air as it passes from the larynx to lips". Vowels are classified on the bases of front, back, high and low manners of articulation and also named respectively (Roach, 2009). According to Skandera and Burleigh (2005) there are three criteria for the description of the vowels i.e. (i) ***the closeness/openness*** criterion refers to the distance between the tongue and the palate and also the position of the lower jaws. (ii) ***Frontness/backness*** refers the case that the body of the tongue pushes forward or backward for the articulation of the front or back vowels respectively if front of the tongue raised high than it will produce central vowel. (iii) ***The shape of the lips*** can be neutral, spread or rounded in the vowel articulation.

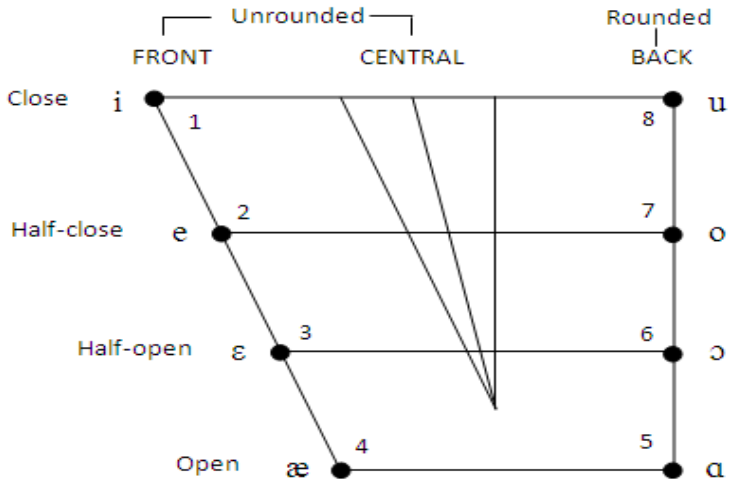


Figure 8: Place & Manner of Articulation in Vowels[6]

2.7. Accent Variation

Accent variation could mark the identification of a person as a member of a specific geographical, socio-economic, ethnical group and also identify his first language.

[6] Retrieved 22-9-13: http://bilingualbambino.blogspot.com/2011/01/variegated-varietals.html

Phonetically, the accent variation is less or more based on the first three formants; F1, F2 and F3 of the vowels (Accents Research, 2004). The lips rounding and the tongue movement are two basic articulatory positions promptly respond to the first three formants (Ladefoged, 2005).

The variation in formant values occur due to the fact that the vocal tract behaves as a non uniform cross sectional tube. This tube naturally resonates at the multiples of 500 Hz approximately. It produced unequally spaced resonance frequencies therefore it is very difficult to associate formants to a consonant's frequency values. But phoneticians agreed that the quality of the vowel could be quantified by measuring the resonant frequencies of the vocal tract (as cited in Keerio, 2010; Parsons, 1987: Reetz & Jongman, 2008).

The accent variation may cause clear difference in the quality of vowels. Vowels are articulated with the vibration of the vocal folds and relatively open vocal tract therefore produced loud acoustic signals. Formant frequencies are the primary acoustic property of the vowels therefore, first three formants (F1 - F3) are critically important among all formant frequencies. They may vary for the same vowel among the speakers of different sex and age groups; change in the position of articulators in the vocal tract causes the change in the formant frequencies. Therefore, formant frequency location is critically important than the position of the articulators. The exact location of the formant frequencies determines the difference between two vowels. *Frontness/backness* and the *height of the tongue* also define the acoustic properties of the vowels. *Vowel height* is inversely correlates with the first formant frequency (F1) i.e. the higher the tongue position, lower will be the formant frequencies and produced higher vowels.

Frontness/backness is reflected in the second formant; F2. If the *lips are more rounded* and the *tongue body moves backward* than the resultant produced F2 will be lower; conversely the higher the F2 value the lips remain either spread or neutral than the more forward is the tongue body in the mouth. The front vowels are related with the difference between F1, F2 means that large difference indicates the presence of front vowels and less difference indicates the back vowels. Formant frequencies above F3 do not provide much information about the quality of the vowel but the formant frequencies of F4 and F5 seem more speaker specific and may provide information about the speaker's identity (Reetz & Jongman, 2008).

2.8. Multilingualism

Multilingualism is a complex field as is obviously clear from the prefix; multi-, is dedicated to the study of comprehension, processing and production of two or more than two languages. Though, in colloquial usage the term "bilingualism" is used as an umbrella term to embody both bilingualism and multilingualism. In its primary stages of exploration, multilingualism is a "rapid growing linguistic area" which deals with the language contacts, linguistic variations, and language changes (Bhatia & William, 2013).

Multilingualism can be explained in many ways but basically, it suggests the ability of an individual to use more than two languages in a situation. Multilingualism can be explained at two levels i.e. the individual level and the societal level. At *the individual level*, multilingualism refers to the competence of a speaker to use two or more languages. At *the societal level* the term multilingualism refers to the use of two or more than languages in a community but it does not mean that all the speakers are equally competent and proficient in more than one language. There are different factors to

influence multilingualism. Communication, education, religion, cultural identity after migration, socio-economic needs for survival in the society and historical or political movements are most important factors causing multilingualism in all over the world (Gorter et al. 2003).

A man is blessed with inevitable senses where listening and speaking abilities are the favors of Allah Almighty on human beings. Of course, the other senses working normally are also the great blessings. At the time of birth, a child is blessed with the language to learn by listening in an environment and explores the whole universe by using listening and speaking abilities. Language is a mean of communication and has its own uses and abuses in our lives. Language has its own moral and ethical values. It is the soul of all trainings and education either belonging to any domain; moral, political, social, spiritual, relating to arts and humanities, economics, science, technology and so on. Mean while we are unable to refuse the importance of mother language in our lives. The phenomenon of the mother language eventually relates with the multilingualism which causes unity (for the speakers of same mother language) and diversity (for the speakers of other mother languages) at a same time. Different psycholinguistic researchers define multilingualism as the use of two or more than two languages. But multilingualism also deals with the cause and effect of one language to the other and may be problematic sometimes (Mughal, 2011).

A human language is a hierarchical process, learned relatively with dynamic changes, a self governing system of meanings which generates syntagmatic and paradigmatic signs. This language system is used to signify and communicate by means of speech communities within the life cycle. This definition captures the important

ideology of a language as a cultural and neurological phenomenon. A language is an obvious matter which does not develop independently outside the cultural perspective (Andrews, Frigau, Voyvodic, Wright, & Voyvodic-Casabo, 2013).

There are 6912 languages in all over the world but most of them are considered as dialect or a variety of a language (as cited in Gorter et al. 2003; Gordon, 2005, www.ethnologue.com). To be a bilingual or multilingual is not an abnormality as supposed by many people but it is a normal and typical need for the majority of population in the world today (as cited in Gorter et al. 2003; Edwards 1994*: 1). The importance of multilingualism at the psycholinguistic level: "The primary question for linguistics should not be Chomsky's (1986) "What constitutes knowledge of language" (p.3), but "What constitutes knowledge of languages?" (as cited in Gorter et al. 2003; Cook 1992: 579).

Actually, languages are not individual entities rather are continuum that vary along a geographical areas and sometimes have no clear boundaries among them. Now a days, linguistic variation and diversity are up raising and important social issues for the languages of the world. The language sustainability is also the major issue to flag for preserving the living languages in the world as the people of all living (endangered) languages are facing lingual threats. Language is the "cultural asset" for establishing relationship between linguistic and economic diversity for the welfare of human beings. The different governments of several countries give official and administrative acknowledgment to only one language which creates the impression that multilingualism is not a widespread phenomenon in the world. But in reality, it would not be easy to find

a monolingual country. In all over the world, the majority of the population can speak more than one language at a time (Gorter, et al., 2003).

Multilingualism is not a compromise but is the crucial reality in a society. The presence of different languages in a country or a state is an attributive quality of the globalization effect. The other idea is the migration of the people, the educational changes, the instructive modifications at the work place, the research and the use of new technologies will also support social adjustments and conscious awareness for the acceptability of the linguistic diversity in a society (Rovira, 2007).

2.8.1. Multilingualism in Pakistan

Pakistan is a multilingual country with six majorly spoken and almost seventy two minor spoken languages where fifty are dialects and there is no unique and independent writing system for dialects (Gordon, 2005). In Pakistan, there are 139 districts having different native languages along with Urdu and English as official languages (Mahmood, 2004). Urdu is the mother tongue of only 7.57 percent of the population (Zia, 2011).

As we are firstly and lastly are Pakistanis despite the fact that we are Punjabi, Sindhi, Balochi, Pakhtun and Saraiki by birth therefore, we are supposed to use Urdu as a national language. But the provincial or local languages are our cultural heritage therefore have right to flourish, develop and play their roles. Along with the national language, provincial languages (Punjabi, Sindhi, Baloch, Pashto and Saraiki) could be used in a provincial assembly by the law. On 23rd March 1956, Urdu and Bengali were declared National Languages of Pakistan (Article no 6, 1956 Act). But it was decided later in the constitution of 1973, article no. 251;

> "Without prejudice to the status of the National language, a Provincial Assembly may by law prescribe measures for the teaching, promotion and use of a provincial language in addition to the National language" (Mahmood, 2004).

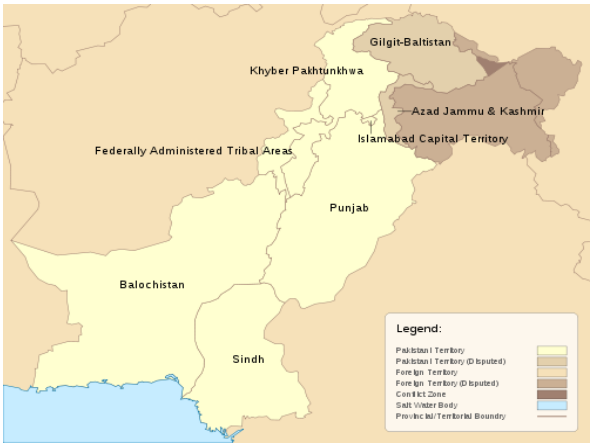


Figure 9: Geographical Division of Pakistan[7]

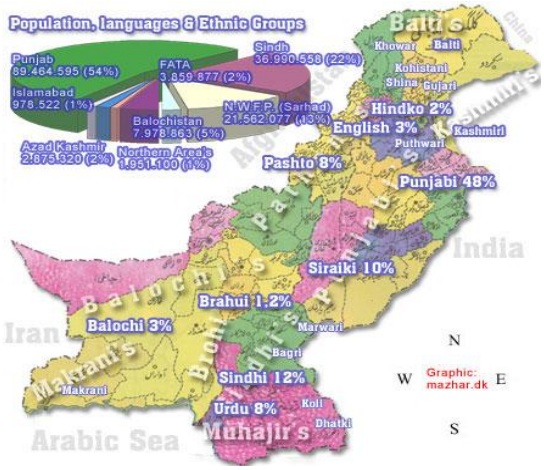


Figure 10: Language Division in Pakistan[8]

2.8.1.1. Multilingualism in Punjab

Punjab is the province of Pakistan, having 43,036,404 number of population[9]. In Punjab, there are almost 36 districts (Zia, 2011) and 30 minorly spoken languages are being used such as; Are, Awankari, Chinavari, Chhachi Dhani, Derawali, Dogri, Ghebi, Gujrati, Haryanvi, Hindko, Jangli, Jandali or Rohi, Jhangochi, Jhangvi, Khatki, Khetrani, Majhi, Malwi, Multani, Pahari, Persian, Pothohari, Raangri, Rajisthani, Rachnavi, Riasati, Saraiki. Shah Puri, Saraiki, Thalochi[10]. While Punjabi is the mother tongue of the people living in Punjab but unfortunately the people of Punjab are very redundant to use

[7] Retrieved 12-8-2013: http://www.googlemaps
[8] Retrieved 12-8-2013: http://www.googlemaps
[9] Retrieved 12-8-2013: http://www.pbs.gov.pk/sites/default/files/population_census/Administrative%20Units.pdf
[10] Retrieved 12-8-2013: http://www.omniglot.com/writing/languages.htm

Punjabi in their social circle therefore, Urdu and English are the official languages and formal languages (Rehman, 2002).

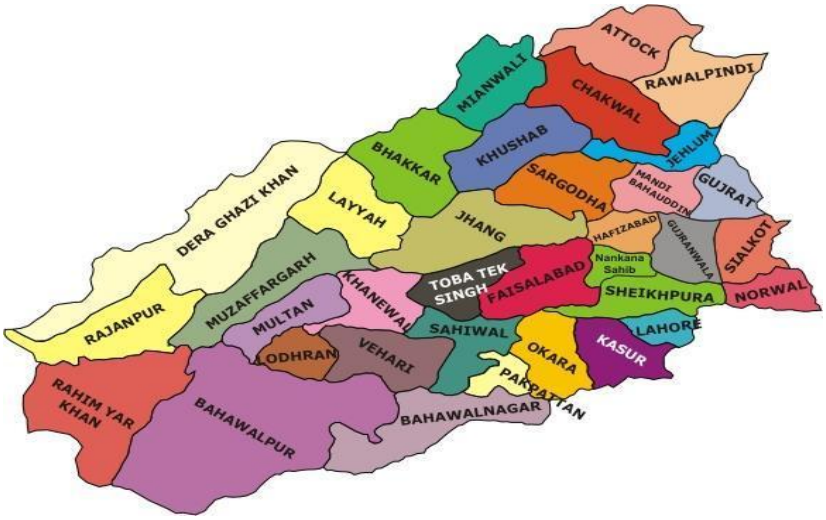


Figure 11: Geographical Division of Punjab[11]

2.8.1.2. Multilingualism in Sindh

According to the census of 1998, in Sind, the estimated population rate[12] was 14,839,862, there are 23 districts (Zia, 2011) and almost 34 minor languages are being used among people such as; Are, Bhaya, Dhatki, Ghera, Goaria, Gujrati, Gurgula, Hasoria, Hazaragi, Jandawera, Katai Meghwar, Kuchi Bill, Kabutra, Kachichi, Koli Parkari, Koli, Kachi, Lasi, Marwari, Matli, Memoni, Mewasi, Nairya, Od, Rabari, Persian, Rardro, Saraiki Sansi, Sindhi Bhil, Tharadari, Vaghri, Wadiyara, Zalavaria Koli[13]. Sindhi is the mother tongue as well as the official language but Urdu is intelligible only among the educated people who are living in the urban areas (Rehman, 2002).

According to the geological and archaeological researches; Sindh is a sea-born land. In 1920s, the discovery of Moen-Jo-Daro (means mounds of dead) revealed that a

[11] Retrieved 12-8-2013: http://www.googlemaps
[12] http://www.pbs.gov.pk/sites/default/files/population_census/Administrative%20Units.pdf
[13] Retrieved 12-8-2013: http://www.omniglot.com/writing/languages.htm

highly developed culture had existed in Sindh (during the era of 2600 and 1900 B.C.) which is as old as Chinese civilization, the ancient Egyptian and Mesopotamian (as cited Keerioayaz, 2010; Advani, 1993; Pithawalla, 1935).

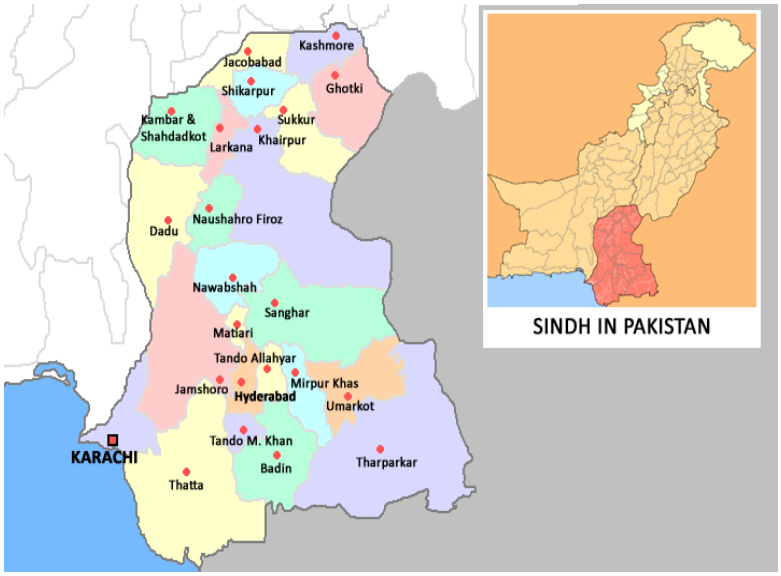


Figure 12: Geographical Division of Sindh[14]

2.8.1.3. Multilingualism in Balochistan

Balochistan is the province of Pakistan with 1,568,780 population size[15]. In Balochistan, there are 30 districts where Balochi is the mother tongue of the population. Most of the people understand and speak Urdu as an official language (Zia, 2011). Despite of the limited population, Balochistan has an uncommon tribal and racial diversity therefore can understand and speak more than two languages (Rehman, 2002). But there are almost 12 minor languages being spoken in Balochistan: Arabic, Brahui,

[14] Retrieved 12-8-2013: http://www.googlemaps
[15] Retrieved 12-8-2013: http://www.statpak.gov.pk/depts/pco/statistics/area_pop/area_pop.html

Dehwari, Hazaragi, Jadgali, Khetrani, Lasi, Pashto, Persian, Punjabi, Saraiki, Sindhi, Waneci, Wanetsi (Pashto dialect)[16].

Figure 13: Geographical Division of Balochistan[17]

2.8.1.4. Multilingualism in Khyber-Pakhtunkhwa

In Khyber-Pakhtunkhwa, the number of districts are 25 (Zia, 2011) and almost 35 minor languages are being used at sub-district level such as; Badeshi, Bashgali, Bateri, Chilisso, Dameli, Dari, Gawar Bait, Gowro, Gujari, Hazaragi, Hindko, Kalkoti, Kalasha, Kalami, Kamviri, Kati, Khowar, Kohistani, Ormuri, Phalora, Persian, Pothohari, Punjabi, Savi, Shina, Saraiki, Turwali, Ushujo, Wakhi, Yidgha but Pashto and Urdu are the majorly spoken languages[18]. As Pashto is the mother tongue and Urdu is an official language of Khyber-Pakhtunkhwa (Rehman, 2002).

[16] Retrieved 12-8-2013: http://www.omniglot.com/writing/languages.htm
[17]Retrieved 12-8-2013: http://www.googlemaps
[18]Retrieved 12-8-2013: http://www.omniglot.com/writing/languages.htm

Figure 14: Geographical Division of Khyber-Pakhtunkhwa (NWFP)[19]

2.8.1.5. Multilingualism in FATA

In FATA, there are 13 districts (Zia, 2011) and almost 9 minor languages are being spoken by the population (Rehman, 2002) i.e. Dogri, Gojri, Kashmiri, Kundal Shahi, Mirpuri, Majhi, Pahari-Potwari and Potwari are being used along with the Urdu[20].

Figure 15: Geographical Division of FATA[21]

2.8.1.6. Multilingualism in Azad Kashmir

In Kashmir, there are 10 (Zia, 2011) districts and almost 9 minor languages[22] are being used for communication (Rehman, 2002) such as Dogri, Gojri, Kashmiri, Kundal

[19] Retrieved 12-8-2013: http://www.googlemaps
[20] Retrieved 12-8-2013: http://www.omniglot.com/writing/languages.htm
[21] Retrieved 12-8-2013: http://www.googlemaps

Shahi, Majhi, Mirpuri, Pahari Potwari, Punjabi and Urdu are also intelligible by the native people.

Figure 16: Geographical Division of Azad Kashmir[23]

2.8.1.7. Multilingualism in Gilgit-Baltistan

In Gilgit-Baltistan, there are 7 districts (Zia, 2011) where people use almost 7 languages[24] (Rehman, 2002) such as Balti, Burushaski, Chainies, Domaaki, Gojal, Shina, Wakhi, along with Urdu as an official language.

Figure 17: Geographical Division of Gilgit-Baltistan[25]

---

[22] Retrieved 12-10-2013: http://www.omniglot.com/writing/languages.htm
[23] Retrieved 12-8-2013: http://www.googlemaps
[24] Retrieved 12-10-2013: http://www.omniglot.com/writing/languages.htm
[25] Retrieved 12-8-2013: http://www.googlemaps

2.9. Six Majorly Spoken Languages in Pakistan

Pakistan is a multilingual country where minor languages are comprehended and used by minority of population therefore are called minor languages and major languages are those understood by majority of population (Gordon, 2005). There are six majorly spoken languages in Pakistan i.e. Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. Brief introduction of these languages are discussed here.

2.9.1. Urdu

According to Ethnologue (2005), Urdu is being spoken by 100 million people in all over the world. It is partially spoken and more than partially understood in South Asian countries i.e. Bangladesh, India and UAE. Urdu is a Turkish word which means *"Camp or Army with its followers"* and major languages participating in the camp of Urdu e.g. Arabic, English, Persian and Portuguese (as cited in Saleem et al., 2012; Saksena).

When the Congress started the movement of 'Hindi Hindustani' Quaid-e-Azam Muhammad Ali Jinnah said: "We know the real object of this scheme is to throttle Urdu". Then, in 1938, Quaid-e-Azam replied to Pandit Jawahar Lal Nehru in answering a question: "Another demand of Muslims is concerning the language and transcription. Urdu is our practically national language. We want constitution protection that Urdu shall not be adversely affected or destructed in any manner." Again in 1941, Hazrat Quaid-e-Azam Muhammad Ali Jinnah said in his address at Aligarh University: "Let me live in Pakistan leading my life in the light of Islamic history and my own culture and according to my own traditions and keeping my Urdu language intact." Quaid-e-Azam Muhammad Ali Jinnah was always in favor of Urdu as a national language of all the Pakistanis,

without any provincial prejudice. During his struggle for Pakistan, in 1946, when the first prime minister, Sir Feroz Khan Noon wanted to deliver his speech in English, Quaid-e-Azam get involved in and asked him to alter his decision for the selection of Urdu in place of English: "The official language of Pakistan shall be Urdu". Quaid-e-Azam Muhammad Ali Jinnah declared Urdu as a national and official language at Dacca Airport in 1948: "The national and official language of Pakistan shall be Urdu alone" (as cited in Mughal, 2011).

In Pakistan, Urdu has been declared as a National language under the constitution of Pakistan in 1973. According to the Article 251 (1) "The National language of Pakistan is Urdu, and arrangements shall be made for its being used for official and other purposes within fifteen years from the commencing day". Urdu and English would be used in writing newspapers, applications, publication of notices and proceedings in courts. The national language, Urdu plays a uniting factor in a democratic society of Pakistan. Being Pakistanis, it is our democratic duty to stick with the *Ideology of Pakistan* by using Urdu in our daily lives. As the purpose behind, *"Nationalizing Urdu*" was to preserve our cultural identity as a nation with our religion, Islam (Mughal, 2011).

Phonetic Inventory

Urdu is phonetically similar to Hindi but is different in alphabetical script and historical characteristics. The pronunciation of Urdu varies with reference to the geographical changes in Pakistan. It is a rich language with a large variety of sounds and all of these sounds can be differentiated on the basis of duration, quality and nasalization (Rehman, 2006).

There is a controversy in the total number of consonants in Urdu language. As according to Hussain (1997) and Raza A. A.(2009) there are thirty six consonants but according to some other studies, there are forty three (Qandeel, Mahmood, & Mahmood, 2012) or forty four consonants (Raza A. A., 2009). The confusion in the number of consonants in Urdu is due to the traced consonants such as; aspirated nasals /mʰ/ and /nʰ / aspirated lateral /lʰ /, aspirated flap /ɽ ʱ/ and aspirated trill /rʰ / (Oxford Urdu English Dictionary, 2013; as cited in Qandeel et al., 2012; Raza A. A., 2009; Saleem et al., 2002). Urdu has Uvular stop sound /q/. It also has borrowed sounds /x/, /ɣ/, /f/ and /z/ which are substituted with /kʰ /, /g/, /pʰ / and /ʤ/ respectively but these sounds are being observed only in the speech of educated people in Pakistan (as cited in Qandeel, 2012; Shackle, 2003).

(As cited in Oxford Urdu English Dictionary, 2013; Saleem et al., 2002)

Table 1: *Urdu Consonantal Inventory*

	Bilabial	**Labio-dental**	**Dental**	**Alveolar**	**Retro flex**	**Palatal**	**Velar**	**Uvular**	**Glottal**
Stop	p b pʰ bʰ		t̪ d̪ t̪ʰ d̪ʰ	ʈ ɖ ʈʰ ɖʰ			k g kʰ gʰ	q	ʔ
Fricative		f v		s z		ʃ ʒ	ɣ	x	h
Affricates				ʧ ʤ ʧʰ ʤʰ					
Nasal	m mʰ		n nʰ				ŋ		
Lateral				l lʰ					
Trill						r rʰ			
Flap					ɽ ɽʰ				
Glide		(w)				j			

Vowels

There are 7 long and 3 short oral and nasal vowels, 3 medial oral vowels (Oxford Urdu English Dictionary, 2013) and a number of diphthongs have also been reported in Urdu language by the Center for Language Engineering, UET, Lahore (www.cle.org.edu.pk.).

Short Vowels

There are 3 short vowels in Urdu language i.e. /ɪ/, /ə/ and /ʊ/ (Oxford Urdu English Dictionary, 2013; as cited in Qandeel et al., 2012; Raza A. A., 2009; Saleem et al., 2002).

Medial (Majhul) Vowels

According to Oxford Urdu English Dictionary (2013) Urdu language also has 3 medial vowels i.e. /e/, /æ/ and /o/. They show mixed properties of short and long vowels means that the medial vowels are audible like long vowels but their duration is larger than short vowels and lesser than long vowels. Most of the time medial vowels are followed by /h/ sound. There is no separate symbol IDs for medial vowels in IPA. Therefore, same symbols are used for long and medial vowels. But for making the difference between long and medial vowels the duration mark (**ː**) is used with long vowels.

Long Vowel

In Urdu, there are 7 long oral vowels i.e. /i:/, /e:/, /æ:/, /a:/, /ɔ:/, /o:/ and /u:/ (Oxford Urdu English Dictionary, 2013; Qandeel et al., 2012; Saleem et al., 2002).

Nasal Vowels

Urdu language has contrastive nasal vowels, equal in number to oral vowels i .e. /ĩ:/, /ẽ:/, /æ̃:/, /ã:/, /ɔ̃:/, /õ:/ and /ũ:/ (Zahid, 2010).

(Quardileteral of Urdu oral and nasal vowels by Raza A. A., 2009)

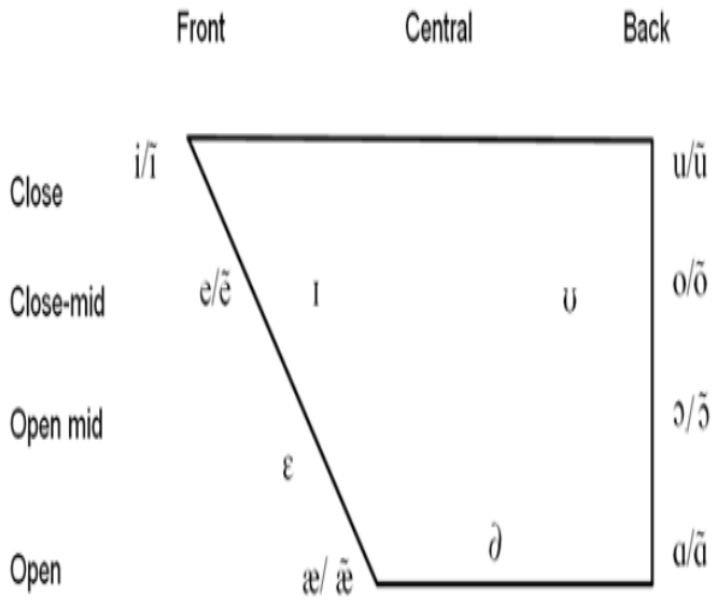


Figure 18: Contrastive Urdu Vowels

List of oral and nasal vowel is given below used in prototype speech annotation at Center for Language Engineering, 2014 (www.cle.org.edu.pk.).

Table 2: *Oral & Nasal Vowels in Urdu*

Sr. no.	**Oral Vowel**	**IPA**	**Nasal Vowel**	**IPA**
1	وُ	u:	وُں	ũ:
2	و	o:	وں	õ:
3	وَ	ɔ:	وَں	ɔ̃:
4	ا،آ	ɑ:	ا‌ں،آں	ɑ̃:
5	ی	i:	یِں	ĩ:
6	ے	e:	یں	ẽ:
7	و	o:	وں	õ:
8	ےَ	æ:	یَں	æ̃:
9	ِ	ɪ	ں ِ	ɪ̃
10	ُ	ʊ	ں ُ	ʊ̃
11	ء،َ	ə	ں َ	ə̃
12	ہ ِ	e		
13	ہ َ	æ		
14	ہ ُ	o		

Syllable Templates in Urdu

There are 8 syllable templates in Urdu (Hussain S. , 2006).

Table 3: *Urdu Syllable Templates*

Sr. no.	Urdu Syllable Template
1	V
2	VC
3	VV
4	VVC
5	CV
6	CVC
7	CVV
8	CVVC

Lexical Stress in Urdu Language

Lexical stress is another important feature of Urdu language. It is an abstract phonological phenomenon, assigned to one or more than one syllable in a word of any language depending on different mechanisms. Two hypotheses are more commonly being used to explain lexical stress in any language i.e. *Hyperarticulation theory*[26] and the *Sonority Expansion Theory*. Acoustic phonological cues for the stressed syllable are F0, quality of vowel and duration of the stressed syllable (consonant and vowel) in a word. In Urdu language, lexical stress is marked on the final heavy syllable but on the other hand

[26] According to Hussain (1997) Hyperarticulation Theory suggests that all phonemically distinctive contrasts are more distinctively articulated and The Sonority Expansion Theory suggests that stress works to increase the sonority feature. Both theories explain some but not all stress related effects for different languages.

it is assuming that final mora[27] in Urdu is extrametrical and therefore not responsible for contributing the weight to the final syllable. A syllable is considered heavy if it has at least two morae. A short vowel and a coda consonant are mono moraic and a long vowel is bi-moraic which contributes a heavy syllable (Hussain S. , 1997).

Summary of Urdu Language

- Urdu is a majorly spoken and national language of Pakistan but accent varies geographically (Saleem, Kabir, Riaz, Rafique, Khalid, & Shahid, 2012).
- It is syllable stress language and it has eight syllable templates (Hussain, 2006).
- It is phonetically similar to Hindi but has different alphabets (Saleem et al., 2012).
- There are 66 phonemes in the phonetic inventory of Urdu:
    - Forty three consonants
    - Seven long oral vowels
    - Three medial oral (Majhul) vowels
    - Three short oral vowels
    - Seven long nasal vowels
    - Three short nasal vowels

[27] According to Hussain (1997) the list of Urdu syllable templates;

Sr. no.	Urdu Syllable Templates	Moraic Weight
1	CV	1
2	CVV	2
3	CVC	2
4	CVVC	3
5	V	1
6	VV	2
7	VC	2
8	VVC	3

2.9.2. Punjabi

Punjabi is an 'Indo-Aryan language' and has a rich cultural heritage. In 11th century, it is recognized as an independent language and is the 10th largely spoken language of the world. It has around 109 million speakers in Pakistan and India who speak Punjabi as their first language which means almost 76 million speakers in Pakistan and 33 million speakers are living in India. In Pakistan, there is no official status of Punjabi language. But in India, it is the first official language in Punjab and in the other provinces is one of an official language along with other 22 languages (Ghai & Singh, 2013).

Punjabi has a large number of speakers but is called a Less Resource Language (LRL) on the account of non availability of the electronic resources. It is also called an inter-continental language (Lata, 2011).

Pakistani spoken Punjabi is more influenced with Perso-Arabic sources and it has different dialects such as Multani, Lehandi and Pothohari, etc. These dialects are different from each other at both phonemic and word level. In India, it is influenced by Sunskrit and is also considered a sister language of Hindi. In Punjabi, ***tone*** is segmental and phonemic in function which causes pitch variations for distinguishing the different meanings of a word that has same vowel and consonant pattern (Karamat, 2012). During the production of a tone, there is neither air frication nor air stoppage occurs in the oral tract which makes Punjabi language phonologically complex, contrastive (Lata, 2011) and unique among other Indo-Aryan languages (Qandeel et al., 2012).

Nasalization is another important but less under stood feature by the Punjabi speakers. The oral sounds are produced with the configuration of oral and nasal tract; the

complete closure of nasal tract produces nasal vowels with the open velopharyngeal port. There are two types of nasalization i.e. *contrastive* nasalization and *contextual* nasalization. The nasalization due to the effect of neighboring nasal sounds is called contextual nasalization and the nasalization produced with the alternation of minimal pairs of oral and nasal vowels is called contrastive nasalization. So, by the change of a nasal vowel with the oral vowel the meaning of the word would be changed (Zahid, 2010).

Stress is not a major feature in Punjabi but helps in distinguishing the disyllabic grammatical words. In Punjabi, all stressed syllables are accented and the accent is the combination of pitch and length therefore, unstressed syllables have lack of pitch and length. ***Intonation*** is the fluctuation in the pattern of pitch as applied to a larger unit than the word. A clause or a sentence could be uttered in more than one way by giving different meaning to the hearer (Lata, 2011).

Phonetic Inventory

In Pakistan, Punjabi has been written in Arabic script and is called Shahmukhi but in India is written in Devangari script and is called Gurumukhi. Punjabi consonants have been divided into six groups (Ghai & Singh, 2013). It is very difficult to find out the exact consonantal inventory of Punjabi language because each author presents a slightly different phonetic inventory by the other. Shahmukhi Punjabi has 31 consonants (Qandeel et al., 2012) and Gurumukhi Punjabi has 35 consonants, 10 vowels, 7 diphthongs and 3 tones (Lata, 2011). The Punjabi stop system of consonants is similar to

Urdu and Hindi but the historical voiced aspirates are replaced with three tones (Bakst, 2012).

There is controversy in the declaration of places and manners of phonemes in Punjabi inventory. According to Lata (2011) and Karamat (2010) the phonemes; /f/ and /v/ are labiodental fricatives but only Qandeel et al. (2012) reported /f/ and /v/ as bilabial fricatives and on the other hand, Lata (2011) has also reported bilabial approximant /w/. There is also a confusion about the phonemes; /ʧ/, /ʧʰ/ and /ʤ/ as palatal stops (Qandeel et al., 2012; Karamat, 2010) or as palatal affricates Lata (2011).

Punjabi also has different borrowed phones (/x/, /ɣ/, /f/ and /z/) from Persian and Arabic (Qandeel et al., 2012; Karamat, 2010). It also has a set of voiceless aspirated stops; /pʰ /, /ʈʰ /, /ʈʰ/ and /kʰ / (as cited in Qandeel et al., 2012; Mahmood, 2011).

Table 4: *Punjabi Consonantal Inventory by Lata, 2011*

	Bilabial	**Labio-dental**	**Dental**	**Alveolar**	**Retroflex**	**Palatal**	**Velar**	**Glottal**
Stops	p pʰ b		ʈ ʈʰ ɖ		ʈ ʈʰ ɖ		k kʰ g	
Fricatives		f v		s z		ʃ	x ɣ	h
Affricates						ʧ ʧʰ ʤ		
Nasals	m			n	ɳ		(ŋ)[28]	
Laterals				l	ɭ			
Flaps				r	ɽ			
Approximant	(w)					j		

[28] Velar nasal /ŋ/ has reported only by Karamat (2010)

Vowels

In Punjabi language, there are 10 vowels and exist in pairs i.e. a long vowel and a short vowel (Ghai & Singh, 2013). Different sources claim different number of Punjabi vowels e.g. the variation in the number of short vowels is; short vowel /ɪ/ and /ʊ/ change with stress into short form of /e/ and /o/ respectively. This important variation of vowels is clearly visible in the following diagram by showing the place of articulation of vowels by Karamat (2010).

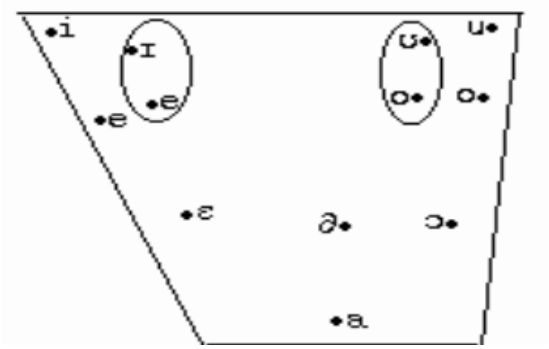


Figure 19: Punjabi Vowels

Table 5: *Punjabi Oral Vowels*

	Ghai & Singh (2013)		**Karamat (2010)**	
	Short Vowels	**Long Vowels**	**Short Vowels**	**Long Vowels**
1	ə	a:	ə	a:
2	ɪ	i:	ɪ	i
3				e
4	e	ɛ:		ɛ
5	ʊ	u:	ʊ	u
6				o
7	o	ɔ:		ɔ

Nasal Vowels

Punjabi language has nasal vowels in contrast of all oral vowels. So, there is larger possibility of *contextual* and *contrastive* nasalization in connected speech. The Punjabi speakers showed relatively greater degree of nasalization for /ə̃/, /ã/, /æ̃:/, /æ̃/, /ĩ/, /ĩ:/ but less for the short vowel /Λ/ and also nasalize both low and high vowels but showed greater nasality for low vowels (Zahid, 2010).

Table 6: *Contrastive Nasal Vowels in Punjabi*

Sr. no.	**Short Nasal vowel**	**Long Nasal Vowel**
1	ə̃	ɑ̃:
2	ĩ	ĩ:
3	æ̃	æ̃:
4	ẽ	ẽ:
5	ʊ̃	ũ:
6	õ	ɔ̃:

Semi Vowels

There are two semi vowels or approximants in Punjabi i.e. /w/ and /j/ (Ghai & Singh, 2013).

Syllable Templates in Punjabi

According to Ghai and Singh (2013), there are 7 templates of Punjabi syllables i.e.

Table 7: *Punjabi Syllable Templates*

Sr. no.	**Punjabi Syllable Templates**
1	V
2	VC
3	CV
4	VCC
5	CVC
6	CVCC
7	CCVC

Summary of Punjabi Language

- According to Karamat (2002), Punjabi is culturally rich Indo-Aryan language since 11th century.
    - Lehandi, Multani, and Pothohari are three spoken dialects of Punjabi in Pakistan.
    - It is a tonal language and its three different tones make it phonemically difficult language.
- It is largely spoken language in Pakistan and India; still called less resource language due to the unavailability of the electronic resources (Lata, 2011).
- It has no official status in Pakistan (Ghai & Singh, 2013).

- Nasalization also makes it difficult language (Zahid, 2010).
- There are fifty eight phonemes in Pakistani Punjabi:
    - Thirty one consonants
    - Ten oral vowels
    - Twelve nasal vowels
    - Seven diphthongs

2.9.3. Sindhi

Sindhi is also a majorly spoken language in Pakistan[29]. In 1996, it is ranked as the 50th largely spoken language of the world[30]. The rate of population for this ranking is outdated now. According to the census of 1981, there were 19 million Sindhi speakers in Pakistan whereas according to the census of 1998, the figure of the population in Sindh was 30.4 million speakers[31].

In Sindh, Sindhi is first language and both Urdu and English are used as second languages in schools and offices. In 2005 Jennifer said, Sindhi language is facing different socio-economic and socio-lingual dangers by Urdu and English languages in Pakistan. Grimes (2001) and David (2003) believed, the other factors such as poor literacy rate (45.29% in 1998 census), social prestige, mass migration, lack of research and developmental resources are becoming the cause of hesitation among the native speakers, shifting to the other prestigious and more prominent languages[32]. All these factors are the cause of reduction in the population of native Sindhi speakers. Others factors are potential threats including the reduction of speakers, less prestige, increase in

[29] Retrieved 12-10-2013: http://www.omniglot.com/writing/languages.htm
[30] Retrieved 14-11-2013http://www.davidpbrown.co.uk/help/top-100-languages-by-population.html,http://www.photius.com/rankings/languages2.html
[31] Retrieved 14-11-2013: http://www.ststpak.gov.pk/
[32] Retrieved 14-11-2013: http://www.sindh.gov.pk/aboutsindh.htm

the rate of urbanization where parents are hesitant to impart Sindhi as mother tongue to their children. The other most important factor is that Sindhi language is difficult to learn because of the large number of phonetic elements (Keerio, 2010).

There is very limited work done in Sindhi linguistics due to migration after the partition of the sub-continent. Before partition, Grierson's work (1919) titled as *Linguistic survey of India*, Ernest Trumpp's work *Grammar of the Sindhi Language* and the work of Sir John H. Marshall titled; *Mohenjo-Daro and the Indus Civilization*, provides a solid historical and linguistic background of Sindhi language along with the geographical discussion but lacks the information of acoustic phonetics (as cited in Keerio, 2010).

The notable work in the acoustic analysis of the voiced implosives[33] of Sindhi has been done by Raza et al. (2004). The study on the segmentation of Sindhi language was carried out by Khawaja et al. (2007) on tracking the dynamics of formants but he has lacking the information about how the segmentation was achieved for the phoneme classes such as the vowels, diphthongs, liquids and glides etc. Other areas in Sindhi language are articulatory phonetics, grammar, orthography and morphological structure refer to the work of authors in this respect (as cited in Keerio, 2010: Trumpp, 1872; Grierson, 1919; Advani, 1993; Jatoi, 1996; Jennifer, 2006).

The research in the field of acoustic phonetics of Sindhi language was ignored a lot. But the work of Keerio (2010) deals with the acoustic phonetic characteristics of the language elemental sounds, accent specific acoustic variations, compilation of

[33] An implosive consonant is produced with the lowering glottis (as down word movement of the piston by creating a short rarification of the air in mouth. When the stricture in mouth released the air moves into mouth (Webster's New World College Dictionary, 2010)

meaningful phonemes with great emphasis on the unique phonetic and linguistic characteristics and easy to learn sound system of Sindhi language.

Phonemic Inventories

Sindhi language uses two writing scripts; Urdu and Hindi which is considered one language with two scripts[34]. While the modern Sindhi uses the *Arabic Nask Script* (among Muslim Sindhi speaker areas) and *Devanagri* script (among Hindu Sindhi speaker areas). But the speaking style of Muslims and Hindus is similar to each other (as cited Keerio, 2010; Jennifer, 2006: Khawaja, 2004).

There are six dialects in spoken Sindhi language in Sindh, Pakistan and name of them are; Thareli, Lasi, Lari, Vicholi (Middle dialect), Utradi and Kachhi. Vicholi or Middle dialect is taken as a standard and is also discussed with reference to linguistic similarities and diversities in detail. Vicholi is important because it shares geographical boundaries with all the other Sindhi dialects. Sanghar district is associated with Vicholi dialect although the dialect of this region is slightly different from middle dialect, Vicholi (Keerio, 2010).

In Sindhi language, there are six consonant classes; stops, implosives, nasals, fricatives, liquids and flaps. Monophothongs, diphthongs and glide consonants are also present. It has forty one consonants and ten vowels. Its phonemic inventory is consisted of the large number of stop consonants ; /p/, /b/, /pʰ /, /bʰ /, /ʤ/, /ʤʱ/, /ʧ/, /ʧʱ/. There are very limited languages in the world which have implosive stops; among them Sindhi, Swahili and Hausa are included. It has contrastive implosive stops; /ɓ/, /ɠ/, /ʃ/, /ɗ/. Sindhi implosives are produced with vibrating vocal cords; therefore are referred as voiced

[34] Retrieved 21-11-13: http://www.indianchild.com/hindi.htm

consonants (Keerio, 2010). These are completely non-aspirated glottal ingressive sounds which are produced by lowering the larynx with vibrating vocal cords (as cited in Keerio, 2010: Jennifer, 2006; Raza, 2004; Nihalani, 1986). In Sindhi language, there are four affricate consonants; /ʧ/, /ʧʰ/, /ʤ/, /ʤʰ/ (as cited Keerio, 2010: Jennifer 2006; Raza, 2004). But there is controversy regarding these consonants in Sindhi language because few authors have classified affricates as stops in Sindhi language (as cited in Keerio, 2010: Mirza 2006; Ladefoged, 1996; Jatoi, 1996).

(Consonantal phonetic inventory of Sindhi language by Keerio, 2010)

Table 8: *Sindhi Consonantal Inventory*

	Bilabial	**Dental**	**Alveolar**	**Retro-flex**	**Post-alveolar**	**Palatal**	**Velar**	**Uvular**	**Glottal**
Plosives	p b pʰ bʰ ɓ	t d tʰ dʰ					k g kʰ gʰ ɠ		
Fricative		f v	s z		ʧ ʤ ʧʰ ʤʰ ʃ				h ɦ
Affricate				ʈ ɖ ʈʰ ɖʰ ɖ͡ʂ	ʃ	j		x ɣ	
Nasals	m		n	ɳ		ɲ	ŋ		
Lateral				l					
Trill				r					

According to Keerio (2010), in Sindhi language, there are five nasal phonemes, bilabial /m/, alveolar /n/, velar /ŋ/, retroflex /ɳ/, palatal /ɲ/. It also has two liquid sounds; lateral /l/ and trill /r/. Sindhi trill and lateral class of consonants are jointly referred as liquids. Sindhi liquids have special features i.e. pronounced voiced at the beginning and middle of the words while voiceless at the final position of the words. There are nine fricative sounds; two labiodental; voiced /v/ and unvoiced /f/, glottal voiced /h+/ and unvoiced /h-/, voiced velar /ɣ/ and voiceless /x/, voiced alveolar /z/ and unvoiced /s/, palatal voiced /j/ and palato-alveolar voiceless /ʃ/.

There are nine redundant alphabet symbols in Sindhi phonemic inventory. Most of these phonemes are borrowed from Persian and Arabic languages are; /s/ ث ص س /z/ ذ ز ظ /h/ ح /t/ ط /k/ ق /ʔ/ ع. There are three pairs of consonants which create confusion in general conversation for the listeners in identifying the pronounced sound. The sounds of confused consonants are /b/ ب, as /bʰ / پ the sound /**ɖ**/, ڈ as /**ɖ̣**/, ڊ and the sound /g/, گ as /gʰ / گھ and vice versa. There are twelve sounds which are not yet part of language alphabets in Sindhi language. They are: (i) /ʔ/, ء (ii) /ʈr/, ٹ ر (iii) /ɖr/, ڈ ر (iv) /lɑ/, لا (v) /ʈʰ / ڈھ (vi) /rʰ /, رھ (vii) /lʰ /, لھ (viii) /nʰ /, نھ (ix) /mʰ /, مھ (x) /ɳʰ/, ٹھ (xi) /m̀/ مين (xii) /yɑ/ 'ى.

In Sindhi language, there are three pairs of interchanging consonants which are used by the speakers of different dialects in communication. For example the speakers of the middle dialect pronounce the word fish as: /məʧʰi/, مڇي with /ʧʰ/, چ sound and the speakers of the Utradi dialect change this sound with sound /ʃ/, ش and pronounce the word fish as: /məʃi/, and in similar way, and the sound /ʤ/, is changed with /z/, ز and the sound /r/, ر is changed with /ʈ/, ڙ and vice versa. There are three pairs of substituting

consonants: (i) the sound /x/ خ is substituted with /k^h/ ک (ii) the sound /f/ ف is substituted with /p^h/ ڦ (iii) the sound /g/ گ is substituted with /ɣ/ غ .

Vowels

According to Keerio, (2010) there are ten vowels in Sindhi language. He has identified the vowels by their minimal pair words means if two words are different by having the difference of only one vowel. But both vowels are produced at identical position then both of these words will have different meanings.

Table 9: *Sindhi Vowels*

	IPA Symbol	**Sindhi**
1	I ɪ	اِي اِ
2	e ɛ	اي اَي
3	ə ɑ	اَ آ
4	ɔ o	اَو او
5	ʊ u	اُ اُو

Syllable Templates in Sindhi

There are 8 syllable templates in Sindhi language (Keerio, 2010).

C Consonant

V Vowel

$\bar{V}$ Long Vowel

Table 10: *Sindhi Syllable Templates*

Sr. no.	Syllable Templates in Sindhi
1	V̄
2	CV
3	CV̄
4	CV̄V
5	CCV̄
6	CCVC
7	CV̄C
8	CVCC

Summary of Sindhi Language

- According to Keerio (2010), Sindhi language is 6th majorly spoken language in Pakistan.
- It has two written scripts and six dialects; Kachhi, Lasi, Lari, Thareli, Vicholi and Utradi
- Vicholi is the standard dialect of Sindhi.
- It has eight syllable template
- It has fifty two phonemes excluding diphthongs;
    - Forty two consonants
    - Ten vowels

2.9.4. Balochi

Balochi is a North West Iranian language and is spoken by almost seven million people (Sabir, 2004). Mostly Balochi speakers belong to Pakistan, Iran, Afghanistan, India, the Arab Gulf States, Turkmenistan, East Africa and also by a number of speakers in North America, Europe and Australia. Balochi language is spoken in South Western Pakistan; Balochistan and small number of population in Punjab and Sindh especially in Karachi. Among other West Iranian languages, Balochi is a controversial language (Ahangar, Oostendrop, & Soohani, 2013). Because it is considered closer to Parthian rather than middle Persian (as cited in Ahangar et al., 2013; Elfenbein 1989). In 2003, Paul declared similarity among Balochi and South Iranian languages and Korn also reported the relationship among Balochi and Iranian languages in the form of a family tree (as cited in Ahangar et al., 2013).

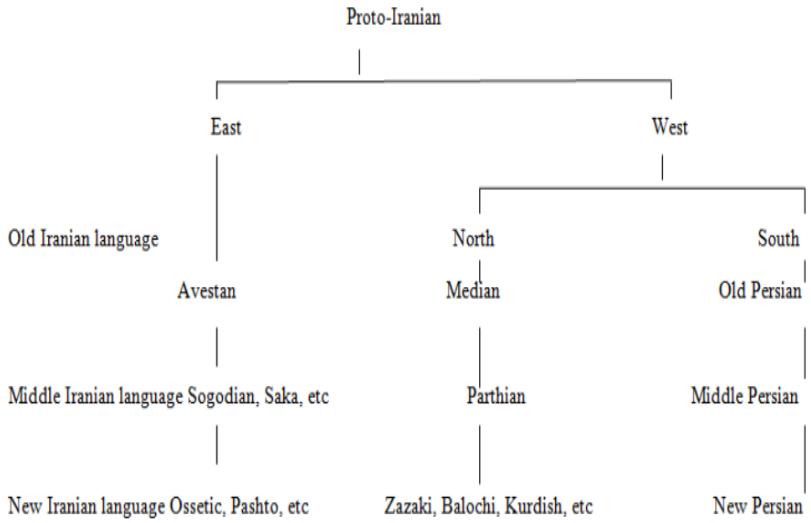


Figure 20: Family Tree of Iranian Languages

According to Sabir (2011), around the world, almost six million people use the *Western* dialect as the standard form of Balochi language. The Central Swarani Balochi (CSB) is a western dialect and has been taken as the standard form of language. It was

also a notion that Balochi was more widely spoken in 19th and 20th century especially in Punjab and Sindh. But nowadays people who recognize themselves as Balochi are going to switch into other languages like Punjabi, Sindhi, Saraiki etc.

In the number of countries, Balochi is used as a medium of speech but not adopted as an official language. In Pakistan, Balochi language was used only as a spoken language till post colonial period but was used as the medium of instruction at the primary level in Balochistan by 1989 to 1992 during the reign of Nawab Akbar Khan Bugti.

Phonetic Inventory

Balochi language has three main dialects; Western Balochi (WBal.), Southern Baochi (SBal.) and Eastern Balochi (EBal.) which is very broad dialects' division because of this reason, most of the dialects are not easily fit into these groups (Ahangar et al., 2013).

Balochi alphabets were developed during 19th century with Roman alphabets devised by Major E. Mockler in 1877. After the independence of Pakistan in 1947, Baloch scholars adopted Perso-Arabic script in writing. Its phonetic inventory is consisted of 25 consonants, 8 vowels and two diphthongs /ai/ and /au/ and the retroflex property of Balochi language is influenced by Indo-Aryan languages (Sabir, 2011).

The uneducated Balochi speakers replaced /f/, /ġ/, /q/ and /x/ with /p/, /g/, /k/ and /h/ respectively and /f/, /ġ/, /q/ and /x/ are found in Arabic and Persian loan words. But the educated population only used Persian pronunciation. Fricative /ɣ/ has been observed in educated people's speech of loan words (Ahangar et al., 2013).

The consonantal inventory of Balochi language (as cited in Ahangar et al., 2013; Baranzehi, 2003) is given below.

Table 11: *Balochi Consonantal Inventory*

	Labial	**Dental Alveolar**	**Retroflex**	**Prepalatal Palatal**	**Velar**	**Glottal**
Plosive	p b	t d	ṭ ḍ		k g	
Affricate				č ǰ		
Fricative	(f)	s z		š ž	(x) (ġ)	h
Nasal	m	n				
Lateral		l				
Flap		r	ṛ			
Glide	w			y		

Vowels

There are 5 long vowels /a̅/, /i̅/, /u̅/, /e̅/, /o̅/ and 3 short vowels /*a*/, /e/, /o/ (as cited in Ahangar et al., 2013; Baranzehi, 2003). According to Soohani, in 2003, the number of short vowels is more than the long vowels. There are 4 long vowels /ɑ:/, /i:/, /o:/, /u:/ and five short vowels /ɑ/, /e/, /o/, /æ/, /ə/ (as cited in Ahangar et al., 2013).

Table 12: *Balochi Vowels*

Sr. No.	**According to Soohani (2003)**		**Sr. no.**	**According to Baranzehi (2003)**	
	Long Vowels	**Short Vowels**		**Long Vowels**	**Short Vowels**
1	ɑ:	ɑ	**1**	a̲	*a*
2	i:	e	**2**	i̲	e
3	o:	o	**3**	u̲	o
4	u:	æ	**4**	e̲	
5		ə	**5**	o̲	

In 2010, Soohani reintroduces the vowel inventory which is similar to the previous work (2003) except few modifications i.e. she does not consider /ə/ as a vowel and considered long /i:/ as a short vowel (as cited in Ahangar et al., 2013).

Syllable Templates in Balochi language

There are 6 syllable templates in Balochi Language (as cited in Ahangar et al., 2013; Soohani 2003, 2010).

Table 13: *Balochi Syllable Templates*

Sr. No.	**Syllable Templates in Balochi**
1	CV
2	CVC
3	CVCC
4	CCV
5	CCVC
6	CCVCC

Summary of Balochi Language

- According to Sabir (2010) Balochi language has three main dialects; Western Balochi, Southern and Eastern Balochi and Soohani (2003, 2010) also informed about six syllable templates (as cited in Ahangar et al., 2013).
- According to Sabir (2011), Balochi is a North West Iranian language and has seven million speakers all over the world.
    - The native Balochi speakers are diverging to other languages because of socio economic issues.
    - It has no official status in Pakistan but was the medium of instruction during the reign of Nawab Akbar Khan Bugti in Balochistan.
    - There are thirty five phonemes in Balochi language;
        - Twenty five consonants
        - Nine vowels
        - Two diphthongs

2.9.5. Pashto

Pashto is an East Indo-Iranian language and there are twenty million Pashto speakers in Pakistan and Afghanistan (Robert, 2000). In 1979, according to a conservative estimation, the total number of speakers in Pakistan is 12 million (including refugees) and 8 million speakers in Afghanistan (Elfenbein, 1997). According to some other researchers, there are forty to fifty thousand people speak Pashto in Pakistan, Iran and Afghanistan. In Pakistan, Pashto is a provincial language in Khyber-Pakhtunkhwa and is also taught till secondary level. In some areas of Balochistan, it is also used as a

spoken language. But it is a native and official language in Afghanistan besides Dari Persian (Din & Rahman, 2011).

Phonetic Inventory

Pashto shares unique property among Iranian languages for having a number of retroflex consonants just like Urdu and Punjabi. This is also true that Urdu and Punjabi are neighboring languages but distantly related with Pashto (Robert, 2000).

According to Din and Rehman (2011), the different dialects and various authors of Pashto language have introduced different number of phonemes. There are five main dialects in Pashto; Yusuzia (North-Eastern dialect), Kandahari dialect (South-Western dialect), Middle Tribal dialect, Central dialect (North-Western dialect), dialect of Quetta (South-Eastern dialect).

According to Elfenbein (1997), it is difficult to describe Pashto phonology. As, Pashto dialects are complexly ramified therefore one has not enough right data to differentiate one dialect from the other dialect. There are various ways in which its dialects can be classified but none of them is entirely satisfactory because of dialectal isoglosses and geographical variations. Therefore, dialects are divided into four classes on the basis of phonology (especially based on the pronunciation of four Pashto consonants i.e. x$^{\check{}}$, g$^{\check{}}$, c, and j).

Those four dialects are differentiated as;

Northwest Group	**Southwest Group**	**Southeast group**	**Northeast Group**
Semi-hard Group)	**Soft Group**		**Hard Group**
xˇ is pronounced ʃ	xˇ is pronounced sˌˇ	xˇ is pronounced ʃ	xˇ is pronounced x/kʰ
gˇ is pronounced ɣˇ/zˇ	gˇ is pronounced zˌˇ	gˇ is pronounced ʒ	gˇ is pronounced /g/
/c/ is pronounced /s/	/c/ is pronounced /ʧ/	/c/ is pronounced /ʧ/	/c/ is pronounced /s/
/j/ is pronounced /z/	/j/ is pronounced /ʤ/	/j/ is pronounced /ʤ/	/j/ is pronounced /z/

In Pashto language, there is a controversy in the number of consonants; 26 (Din & Rehman, 2011) or 30 (Robert, 2000) and the cause of difference between phonetic inventories; may be the ‘nativeness effect of the language. Yusafzai dialect of Pashto language is taken as a standard dialect. There is controversy in the presence of fricatives in Pashto language. All the phonemic information of Pashto language shows that labiodental fricatives e.g. /f/, /v/, /θ/, /ʒ/ and /ð/ are not existing (Din & Rehman, 2011) but Robert (2000) reported the labial fricative /f/.

According to Din and Rehman (2011), in Pashto /w/ is glide without lip roundedness. Labiodental voiced fricative has been observed as a noisy bilabial /w/. The dental voiceless fricative has been observed as dental voiceless stop. There is no labial and dental fricatives but these sounds are pronounced similar to the dental stops found in Southern Irish English. The dental voiced fricative has been pronounced as dental voiceless stop by some of the native speakers because of the orthography of the sounds written with ‘th’ which are pronounced as dental voiceless stop. The post alveolar voiced

fricative is pronounced like a post alveolar voiced affricate of English. /ʒ/ is pronounced as velar voiced stop because of the presence of /ʤ/ in phonemic inventory.

Pashto language has retroflex nasal sound /ɳ/ but has no labial and labiodental fricatives (Ijaz, 2002-2003). The major difference between Southern dialect (spoken in Balochistan, Waziristan and Banu areas) and Yusufzai dialect (spoken in Peshawar) is; /ʧ/ and /ʤ/ are consistently present in Southern dialect while in Yusufzai dialect are pronounced as /s/ and /z/ (Elfenbein, 1997).

Table 14: *Pashto Consonantal Inventory (Elfenbein, 1997)*

	Bilabial	**Dental**	**Palato-Alveolar**	**Retroflex**[35]	**Velar**	**Uvular**	**Glottal**
Plosives	p b	t̪ d̪		ʈ ɖ	k g	(q)	(ʔ)[36]
Fricative	(f)[37]				x ɣ[38]		h
Affricate		c j	cˇ jˇ				
Spirants		s z	sˇ zˇ	xˇ gˇ			
Nasals	m	n		ɳ	ŋ		
Lateral		l					
Flap		r		ɽ			
Glides	w		y				

Stress is the most important element in Pashto language. But the phonetic characteristics, coalescence and controversial patterns of vowel harmony affect the stress

[35] all retroflex consonants were presented by doubling the IPA symbols /ʈ/ as tt and /ɖ/ as dd (Robert, 2000)
[36] present on in Arabic loan words
[37] /f/ is pronounced as /p/ by Robert, 2000
[38] /x/ is presented with the alternative symbol /kʰ/ and /ɣ/ with /gʰ/ (Robert, 2000)

a lot. Each word has a primary stress which is differently determined for each word but most probably stress falls on the final syllable, first or penultimate syllable of the word (Robert, 2000). So, lexical items can be distinguished on the bases of stress. There are three types of stress i.e. weak, medium and strong. The Pashto speakers use lexical stress to differentiate homograph for getting different meanings (Ijaz, 2002-2003). Stress is also used to change number, tense and gender in Pashto language (as cited in Ijaz, 2002-2003; Comrie, 1987).

Vowels

In Pashto language, according to Din and Rehman, (2011) there are nine vowels but Robert (2000) has claimed seven vowels and according to Ijaz (2002-2003) there are seven short and three long vowels. She said, Pashto speakers randomly use short vowels but the long vowels are only used for pronouncing loan words of Persian and Arabic languages.

In Pashto, the long front close vowel /i/ has a short counterpart /ɪ/ vowel and long back high /u/ vowel has a short counterpart /ʊ/ vowel. The front and back vowels are asymmetrical i.e. there are only two back /u/ and /o/ and two front vowel /i/ and /æ/ vowels. Due to the lesser number of vowels the articulatory space between vowels is very limited (Din & Rehman, 2011).

Due the vowel harmony process, the mid vowel /o/ and /ee/ converted into /u/ and /i/ by raising only when the following syllables contain high vowels and the middle short vowel schwa /e/ converts into /i/ and /u/ vowels if followed by rounded consonants (Robert, 2000).

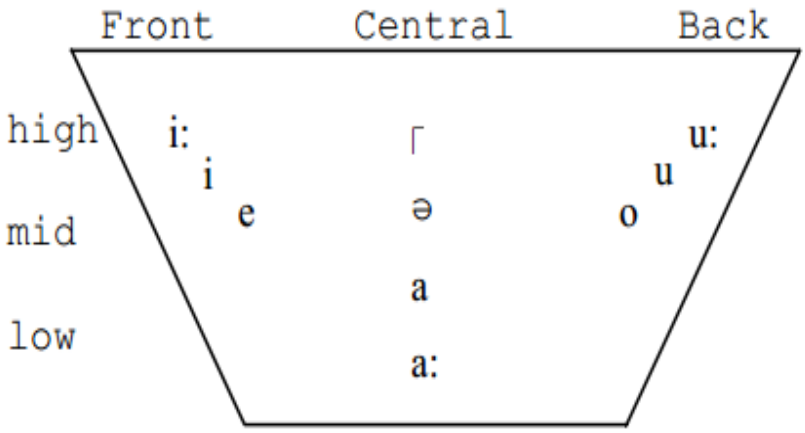


Figure 21: Pashto Vowels (by Ijaz, 2002-2003)

Table 15: *Pashto Vowels*

	Pashto Vowels by Din and Rehman, 2011	**Pashto Vowels by Robert T., 2000**	**Pashto Vowels by Ijaz M.**
1	I	i	i
2	ɪ		ɪ
3	e	ee	e
4	Ə	e	ə
5	u	u	u:
6	ʊ		u
7	æ		
8	a	aa	a:
9	o	o	o
10		a	a

Summary of Pashto Language

- According to Din and Rehman (2011), Pashto is an East Indo Iranian language and has twenty million speakers in Pakistan, Iran and Afghanistan.
    - It is a provincial language and also the medium of the instruction in Khyber-Pakhtunkhwa.
    - It has five main dialects; Yusuzia, Kandahari Dialect, Middle Tribal Dialect, Central Dialect and the Dialect of Quetta
- According to Robert (2000), It has forty two phonemes;
    - Thirty three consonants
    - Nine vowels

2.9.6. Saraiki

Saraiki is called the sister language of Sindhi because Saraiki speaking area was remained the area of Sindh and the spoken dialect of Sindhi language was called "Siroki" means northern. Later in 8th century, Saraiki speaking area get separated from Sindh and was called Multan (as cited in Abbas, 2012; Mehr, 1967). In 1818, the ruler of Punjab, Ranjeet Singh attacked and occupied Multan by adding it into Punjab. Punjabi influenced a lot to Saraiki and got mixed with it. Which became the cause of controversy over the isoglosses of both languages i.e. Punjabi linguist claimed that Saraiki is the dialect of Punjabi language and Saraiki linguist claimed that Saraiki is an individual and independent language. But most of the Western linguists accepted it as an independent language (as cited in Abbas, 2012: Jukes, 1900; Grierson 1919; Smirnov, 1975; Shackle 1972, 1976, 2007a, 2007b). The given picture shows the formation of the Saraiki language: (Abbas, 2012).

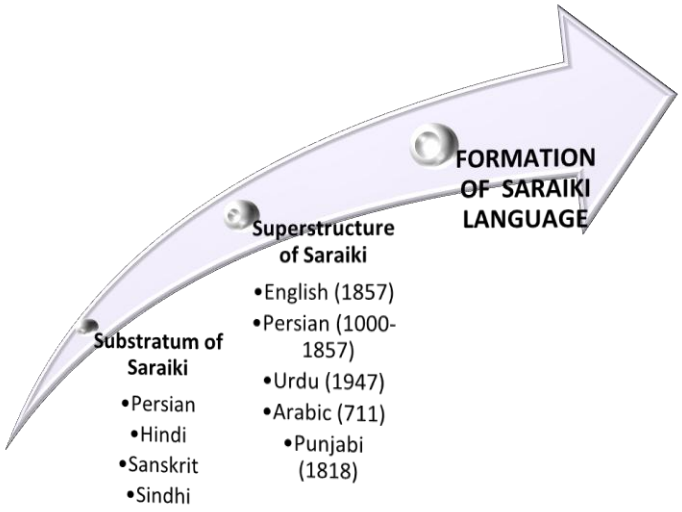


Figure 22: Formation of Saraiki Language

In 1962, the word 'Saraiki' is firstly used in the meeting held under the banner of *Bazm-e-Saqafat,* Multan and this variety is sometimes called *Multani* because of the historical significance of Multan. In 1980, Rasoolpuri calls this a 'pure' form of Saraiki (Awan, Baseer, & Sheeraz, 2012). The meaning of the word Saraiki has many interpretations by different researchers. The most important is that Saraiki is extracted from the word *'Asurki'* which is extracted from the Saraiki word 'Surya' meaning the sun and Asury was the nation who worshipped the sun and ruled in Multan (as cited in Latif, 2006; Ahsan, 1990).

Saraiki belongs to Indo Aryan family and is the language of inhabitants of Indus Valley (Latif, 2006). Since it belongs to the languages of the North Western zone, its dialects bent into Punjabi and Sindhi by showing 85% lexical similarity to Sindhi and 68% with Sansi, Odki and Dhatki (as cited in Latif, 2006; Baart, 2001). Saraiki is an archaic language, found in the Indus civilization i.e. Ddedawar, Harappa, Kot Deji, Multan, Mohenjodaro and Texila (as cited in Latif, 2006; Ahsan 1990). It got the status of

an independent language during the reign of General Zia-ul-Haq in 1981 (as cited in Latif, 2006; Feroz).

Saraiki is a regional language (as cited in Latif, 2006; Gardezi, 1996) and widely spoken in central areas of Pakistan; South Punjab, some parts of Balochistan, Sindh and Khyber-Pakhtunkhwa in Pakistan (Awan et al, 2011). But it is facing an economical competition with dominant languages; Urdu and English. The population census of 1981 treated it as an independent language, spoken by 9.83% of the total population in Pakistan (Rehman T. , 2002). But Wagha (1990) claims that due to some social, political and economical reasons this figure is underestimated. According to Paul (2009) 13,843,106 people speak Saraiki in Pakistan and 20,000 people in India. Haq (1967) claims, in Pakistan the area where Saraiki is being spoken is 48093 sq. miles. It is a spoken language and most of the speakers are unaware of the Saraiki script although it has ancient roots but has no literature (as cited in Latif, 2006).

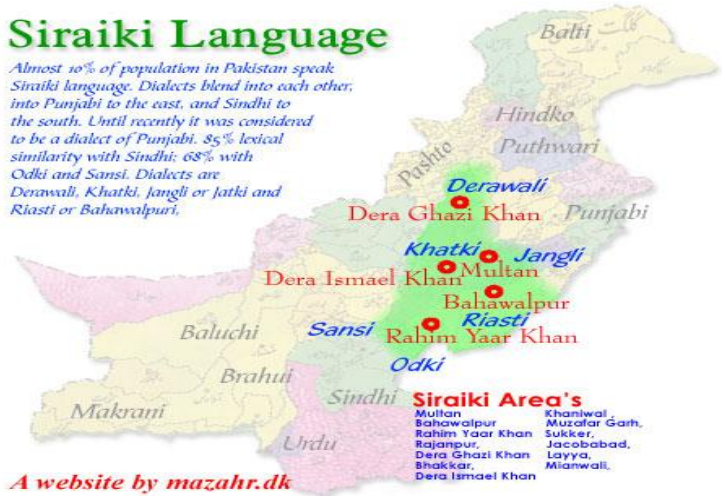


Figure 23: Distribution of Saraiki Language in Pakistan[39]

[39] Retrieved 12-08-13: http://www.googlemaps

Saraiki has attracted the attention of foreign researchers including orientalists like George Grierson, E. O. Brian, Christopher Shackel, Trump, Wilson and Jukes. But almost all linguistic works deal with the historical debate of the language. Most of the researches were carried out by the non native people; Mehr (1967) made the first serious attempt to study it. The efforts of Zami (1963, 1970, 1982), Fikri (1970), Mughal (1990, 1996, 2000, 2003), Wagha (1990) were all focused on the study of history, vocabulary and grammar of the language. Since the emergence of Pakistan, the research in the field of Saraiki language has been very slow (Nasir, 2012).

It is considered the mother tongue or the first language for the people of the central Pakistan whereas the second language for the rest of Pakistani population. No other Pakistani language is spoken as frequently as Saraiki is used. In 1976, Rasoolpuri reinforces this idea by accepting that Saraiki language is the only Pakistani language which is spoken and comprehended by all the provinces in Pakistan (Haq, 1967).

Phonetic Inventory

According to Awan et al (2012), Saraiki has six verities or dialects; Jhangi, Shahpuri, Southern variety, Sindhi variety, Central variety and Northern variety. These varieties are divided on the basis of regions they belong to; Multani, Riasti (Bahawalpuri), Dera Wali, Jhangi, etc. However, all these varieties are given an agreed and collective name i.e. Saraiki. *Central variety* is considered as a standard dialect among all varieties and the speakers of this variety are found in the districts of Multan, Bahawalpur, Muzaffargarh and also in the Northern parts of D. G. Khan.

The total number of Saraiki phonemes are; 94 and not 48, 56 or 58 believed by Shackle, Rasoolpuri and Haq. The problem with their understanding and description is that they ignore the fact that Saraiki monophthongs and diphthongs are separate sounds. It is a syllable time language and has 40 consonants, 23 vowels and 31 diphthongs.

In Saraiki language, there are sixteen plosives and all of these are separate phonemes, not the allophones because they create different words with distinguishing meanings. /t̪/, /d̪/, /t̪ʰ/, /d̪ʰ/, /n̪/, /s/, /z/ are dental sounds and /ʧ/, /ʤ/, /ʧʱ/, /ʤʱ/ are affricate sounds. It has six retroflex sounds i.e. /ʈ/, /ɖ/, /ʈʱ/, /ɖʱ/, /ɳ/, /ɽ/ and two velar sounds /x/ and /ɣ/. It also has one approximant i.e. /j/. But Saraiki has no /θ/ and /ð/ dental sounds. Saraiki has four implosives /ɓ/, /ɗ/, /ʄ/, /ɠ/ but /l/, /ɾ/ are alveolar, /r/ is flap, /m/ is bilabial, /h/ is glotto-velar and /ŋ/ is nasal sounds.

Table 16: *Saraiki Consonantal Inventory (Awan et al., 2011)*

Place **Manner**	**Bilabial**	**Libio-dental**	**Dental**	**Alveolar**	**Palato Alveolar**	**Retroflex**	**Palatal**	**Velar**	**Glottal**
Plosive	p b pʰ bʰ		t̪ d̪ t̪ʰ d̪ʰ			ʈ ɖ ʈʰ ɖʰ		k g kʰ gʰ	
Affricate					ʧ ʤ ʧʰ ʤʰ				
Nasal	m		n̪			ɳ		ŋ	
Fricative		f v	s z				ʃ	x ɣ	h
Lateral				l					
Tap/Flap				ɾ		ɽ			
Glides							j		
Implosive	ɓ		ɗ		ʄ			ɠ	

Vowels

According to Awan et al (2011), there are 23 monophthongs or vowels in Saraiki. It has five long vowels: /i:/, /a:/, /o:/, /e:/, /u:/ and eight short vowels: /i/, /a/, /e/, /ɛ/, /æ/, /ʌ/, /u/, /ɔ/. It also has four long nasal vowels: /ũ:/, /ĩ:/, /õ:/, /ẽ:/ and six short nasal vowels i.e. /ũ/, /ã/, /õ/, /ẽ/, /æ̃/, /ʌ̃/. The location of Saraiki short monophthongs is shown in figure 24.

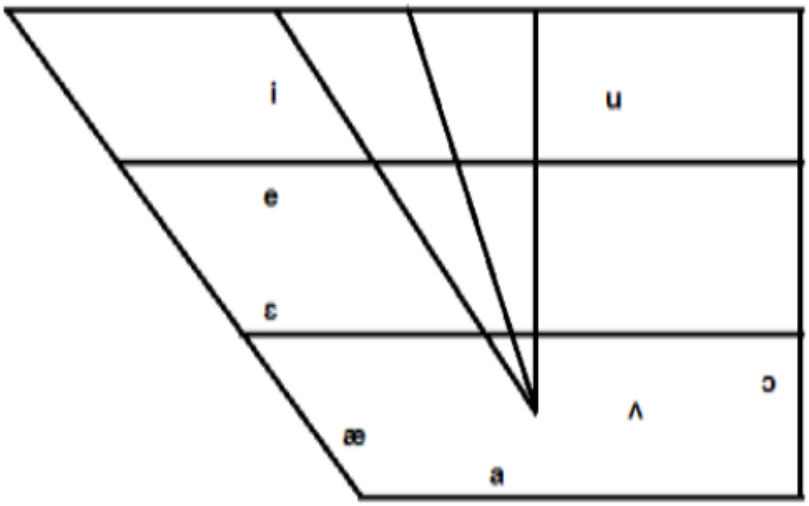


Figure 24: Saraiki Vowels

Table 17: *Saraiki Vowels*

Sr. No.	Oral Vowels	Nasal Vowels
1	ɪ:	ĩ:
2	a:	-
3	o:	õ:
4	e:	ẽ:
5	u:	ũ:
6	i	-
7	a	ã
8	e	ẽ
9	ɛ	-
10	æ:	æ̃
11	ʌ	ʌ̃
12	u	ũ
13	ɔ	õ

Diphthongs

According to Awan et al (2012), Saraiki has 31 diphthongs: /ʌi/, /ɔi/,/ui/, /kʰui/, /ei/, /ae/, /ʌe/, /ie/,/oe/, /ue/, /ao/, /ʌo/, /eo/, /io/, /aɛ/, /au/, /iu/, /ea/, /ua/, /iʌ/, /ʌ /, /aĩ/, /uĩ/, /uẽ/, /iõ/, /aũ /, /iũ/, /eũ/, /ʌũ/, /uã/, /ʤuã/, /eã/, /iã/. It does not have /əʊ/ and /ai/ diphthong and no diphthong ends at /ə/ (because it is syllable time language).

Syllable Templates

According to Awan et al (2012), Saraiki is a syllable time language because all syllables occur at regular interval of time. There is no vowel reduction in Saraiki because of the syllable timing. Saraiki has 10 syllable templates;

Table 18: *Saraiki Syllable Templates*

Sr. no.	**Saraiki Syllable Templates**
1	V:
2	CV
3	CV:
4	CCV
5	VC
6	V:C
7	VCC
8	CVC
9	CCVC
10	CVCC

Summary of Saraiki Language

- Saraiki is an Indo Aryan language and is a regional language in Pakistan (Latif, 2006).
- It recognized as an independent language in the government of General Zia ul Haq (as cited in Latif, 2006; Feroz).

- It has six dialects; Central variety, Jhangi, Shahpuri, Southern variety, Sindhi variety, and Northern variety.
- According to Awan et al (2012), Saraiki is a syllable time language. It has ten syllable templates and has ninety four phonemes;
    - Forty consonants
    - Twenty three vowels
    - Thirty one diphthongs

Comparison of Consonantal Inventories by Using Literature Review

As a first step; comparison of consonantal inventories (of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki) has been done with the help of literature review. The first step is further divided into two steps by comparing the differences or gaps in the places and manners of articulation.

2.10. Step I: Analysis of Phonetic Inventories on the basis of Place of Articulation

According to Skandera and Burleigh (2005) almost all consonantal sounds are produced with the obstruction of the pulmonic air stream in pharynx, vocal or nasal tract. The speech organs that are basically involved in the production of a particular speech sound are most important in terms of place of articulation. A consonantal sound is produced with the contact of the lower, actively mobile speech organs to upper, passively immobile speech organs. ***Labial*** sounds are produced with the lips' movement and form a large group in consonantal inventory. All dental, alveolar, post-alveolar and retroflex sounds are produced with the involvement of the tongue tip therefore are also called ***'apical sounds'.*** The palatal, the velar and the uvular sounds are also called ***'dorsal sounds'*** because all these sounds use body of the tongue. Almost all these consonants are discussed here in detail;

2.10.1. Bilabials

The sounds produced with the contact of lips by creating obstruction of the air stream in the oral tract (Skandera & Burleigh, 2005). /p/, /b/, /m/ are bilabial sounds in Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. /p^h/ and /b^h/ are bilabial aspirated stops found in Urdu , Sindhi and Saraiki languages , Punjabi has only voiceless aspirated stop /p^h/. Bilabial fricative /f/ is only present in Balochi and Pashto languages. Punjabi, Balochi and Pashto have bilabial glide /w/. Only Sindhi and Saraiki have bilabial ingressive stop /ɓ/. Bilabial aspirated /m^h/ is only present in Urdu language but in rare cases.

Table 19: *Comparison of Bilabial Stops among Six Languages*

	Urdu Bilabial	**Punjabi Bilabial**	**Sindhi Bilabial**	**Balochi Labial**	**Pashto Bilabial**	**Saraiki Bilabial**
Stop/ Plosives	p b p^h b^h	p p^h b	p b p^h b^h	p b	p b	p b p^h b^h
Fricative				(f)	(f)	
Nasal	m m^h	m	m	m	m	m
Glide		w		w	w	
Implosive			ɓ			ɓ

2.10.2. Labiodentals

The sounds produced with the obstruction of air stream by the contact of lower lip against the upper teeth are called labiodental sounds (Skandera & Burleigh, 2005). /f/ and /v/ are articulated as labiodental fricative in Urdu, Punjabi and Saraiki. /w/ is labiodental glide in Urdu language.

Table 20: *Comparison of Labiodentals among Six Languages*

	Urdu Labiodental	**Punjabi Labiodental**	**Sindhi Labiodental**	**Balochi Labiodental**	**Pashto Labiodental**	**Saraiki Labiodental**
Stop						
Fricative	f v	f v				f v
Glide	(w)					

2.10.3. Dental/interdentals

Dental sounds are produced with the partial or complete contact of tongue tip with the back of the upper teeth (Skandera & Burleigh, 2005). /t̪/, /d̪/, /t̪ʰ /, /d̪ʰ / are articulated as dental stops in Urdu, Sindhi and Saraiki, /t̪/, /d̪/ are dental stops in Punjabi, Sindhi, Balochi, Pashto but Punjabi has only voiceless dental stop /t̪ʰ/. Balochi and Pashto have no aspirated labiodental stops. Dental implosive /ɗ/ is present only in Saraiki. /n/ is dental nasal in Urdu, Punjabi, Balochi and Saraiki languages. /s/ and /z/ are dental fricative in Balochi and Saraiki but are spirants in Pashto. Aspirated dental nasal /nʰ/ is present only in Urdu language. Lateral phoneme /l/ is dento-alveolar sound in Balochi while dento-latral in Pashto language. /r/ is dento-alveolar flap in Balochi and Pashto.

Table 21: *Comparison of Dentals among Six Languages*

	Urdu Dental	**Punjabi Dental**	**Sindhi Dental**	**Balochi Dental alveolar**	**Pashto Dental**	**Saraiki Dental**
Stop	t̪ d̪ t̪ʰ d̪ʰ	t̪ t̪ʰ d̪	t d tʰ dʰ	t d	t̪ d̪	t̪ d̪ t̪ʰ d̪ʰ
Fricative				s z		s z
Affricates					c j	
Nasal	n nʰ			n		n̪
Spirants					s z	
Lateral				l	l	
Flap				r	r	
Implosive						ɗ

2.10.4. Alveolar

The alveolar sounds produced when the tip of the tongue is coming near or touches the bony alveolar ridge area (Skandera & Burleigh, 2005). /n/ is alveolar voiced nasal in Punjabi and Sindhi. /s/ and /z/ are alveolar fricative in Urdu, Punjabi, Sindhi and Pashto languages. /l/ is alveo-lateral in Urdu and Saraiki, dento-lateral in Punjabi and dento-alveo-lateral in Balochi but is voiced alveolar tap in Pashto language. /r/ is voiced alveolar glide in Pashto but is alveolar flap/tap in Urdu, Punjabi, Balochi and Saraiki languages.

Table 22: *Comparison of Alveolar Consonants among Six Languages*

	Urdu Alveolar	**Punjabi Alveolar**	**Sindhi Alveolar**	**Balochi Alveolar**	**Pashto Alveolar**	**Saraiki Alveolar**
Stops	ʈ ɖ ʈʰ ɖʰ					
Fricatives	s z	s z	s z		s z	
Affricates	ʧ ʤ ʧʰ ʤʰ					
Nasal		n	n			
Lateral	l lʰ	l				l
Tap/Flap		ɾ				ɾ
Glide					r	

2.10.5. Post-alveolar

When the tip of the tongue reaches just behind the alveolar ridge, the post alveolar sounds are produced (Skandera & Burleigh, 2005). /ʧʰ/, /ʤʰ/ are post alveolar aspirated fricative in Sindhi. /ʃ/ is post alveolar affricate in Sindhi, post alveolar affricates in Pashto language. Post alveolar Implosive /ʃ/ is present only in Sindhi and Saraiki languages. In Pashto language, the speakers of Northwest and Northeastern dialects pronounce /sˇ/ and /zˇ/ as fricatives /s/ and /z/ while the speakers of Southwestern and Southeastern groups pronounce as affricates /ʧ/ and /ʤ/ respectively.

Table 23: *Comparison of Post Alveolar Consonants among Six Languages*

	Urdu Post - Alveolar	**Punjabi post-alveolar**	**Sindhi Post-alveolar**	**Balochi Post-alveolar**	**Pashto Palato-Alveolar**	**Saraiki Palato Alveolar**
Fricative			ʧ ʤ ʧʰ ʤʰ		ʃ	
Affricates			ʃ		cˇ jˇ	ʧ ʤ ʧʰ ʤʰ
Spirants					sˇ zˇ	
Implosive			ʃ			ʃ

2.10.6. Retroflex

Retroflex sounds are produced with the curling back of the tongue tip towards the hard palate (Skandera & Burleigh, 2005). /ʈ/ and /ɖ/ are retroflex stops in Urdu, Punjabi, Balochi (with alternative symbols /ṭ/, /ḍ/ respectively) Pashto and Saraiki but in Sindhi are retroflex affricates. /ʈʰ/, /ɖʱ/ are aspirated retroflex stops in Urdu and Saraiki and are retroflex affricates in Sindhi but Punjabi has only voiceless aspirated retroflex stop /ʈʰ/. Retroflex implosive stop /ɗ̢/ is present only in Sindhi language. Retroflex tap/flap /ɽ/ is present in Urdu, Punjabi, Balochi (with alternative symbol /ṛ/) and Saraiki but is retroflex glide in Pashto language. But the aspirated retroflex tap /ɽʱ/ is present only in Urdu but in rare cases. Retroflex nasal /ɳ/ is present in Punjabi, Sindhi, Pashto (with an alternative symbol /ṇ/) and Saraiki. /ɭ/ is retroflex lateral in Punjabi and Sindhi languages. /r/ is retroflex trill in Sindhi language.

In Pashto language, /xˇ/ and /gˇ/ are spirants phonemes means, these sounds are pronounced differently in different dialects. As the speakers of North western group pronounce them as /ʃ/ and /ɣˇ or zˇ/, the speakers of Northeastern group pronounce them as /x or kʰ / and /g/ while the speakers of South western and South Eastern group pronounce these sounds as /ʧ/ and /ʤ/ respectively (Elfenbein, 1997).

Table 24: *Comparison of Retroflex among Six Languages*

	Urdu Retroflex	**Punjabi Retroflex**	**Sindhi Retroflex**	**Balochi Retroflex**	**Pashto Retroflex**	**Saraiki Retroflex**
Stop		ʈ ʈʰ ɖ		ʈ̣ ɖ̣	ʈ ɖ	ʈ ɖ ʈʰ ɖʰ
Affricates			ʈ ɖ ʈʰ ɖʰ ɖʃ			
Spirants					xˇ gˇ	
Nasal		ɳ	ɳ		ɳ̣	ɳ
Lateral		ɭ	l			
Tap/Flap	ɽ ɽʰ	ɽ		ɽ̣		ɽ
Glide					ɽ	
Trill			r			

2.10.7. Palatal

Palate is divided into two parts i.e. soft palate and hard palate. Hard palate is also called "*roof of the mouth*" (Roach, 1998, pp. 9) and the sounds produced with the contact of the tongue body to the hard palate are called palatal sounds. Palato alveolar sounds are produced when the tip of the tongue touches the alveolar ridge and the blade of the tongue is raised towards the hard palate simultaneously (Skandera & Burleigh, 2005).

/ʧ/ and /ʤ/ are palatal affricates in Punjabi and Balochi. /ʧʰ/ is voiceless aspirated palatal affricate in Saraiki. The sound of /ʃ/ is palatal fricative in Urdu, Punjabi, Saraiki and prepalatal palatal fricative in Balochi (with alternative symbol /š/) and /ʒ/ is voiced palatal fricative in Urdu and prepalatal palatal voiced fricative in Balochi (with alternative symbol /ž/). The sound /j/ is palatal glide in Urdu, Pashto and Balochi (with alternative symbol /y/) palatal affricate in Sindhi and palatal approximant in Saraiki but is voiced in Sindhi and Pashto languages. Sindhi language has another voiced palatal nasal /ɲ/.

Table 25: *Comparison of Palatals among Six Languages*

	Urdu Palatal	**Punjabi Palatal**	**Sindhi Palatal**	**Balochi Prepalatal palatal**	**Pashto Palatal**	**Saraiki Palatal**
Fricative	ʃ ʒ	ʃ		š ž		ʃ
Affricates		ʧ ʧʰ ʤ	j	č ǰ		
Nasal			ɲ			
Trill	r r^h					
Approximant	j	j		y	j	j

2.10.8. Velar

Velum or soft palate is an important oral organ which allows the air to pass through the mouth and nose (Roach, 1998, pp. 9). When the back of the tongue comes in contact with the lower side of the velum or soft palate, velar sounds are produced (Roach, 1998, pp. 9; Skandera & Burleigh, 2005).

/j/ is velar glide in Punjabi. /k/, /g/ are velar stops in Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages. /k^h /, /g^h/ are aspirated velar stops in Urdu , Sindhi and Saraiki but Punjabi has only /k^h /voiceless aspirated velar stop. Sindhi and Saraiki also have implosive velar stop /ɠ/. The sounds of /x/ and /ɣ/ are velar fricatives in Punjabi, Balochi (/ɣ/ with alternative symbol /ġ/), Pashto and Saraiki. Velar nasal /ŋ/ is also present in Urdu, Punjabi, Sindhi, Pashto and Saraiki languages.

Table 26: *Comparison of Velars among Six Languages*

	Urdu Velar	**Punjabi Velar**	**Sindhi Velar**	**Balochi Velar**	**Pashto Velar**	**Saraiki Velar**
Stop	k g k^h g^h	k k^h g	k g k^h g^h	k g	k g	k g k^h g^h
Fricative	ɣ	x ɣ		(x ġ)	x ɣ	x ɣ
Nasal	ŋ	(ŋ)[40]	ŋ		ŋ	ŋ
Glide		j				
Implosive			ɠ			ɠ

[40] Velar nasal /ŋ/ has reported only by Karamat (2010)

2.10.9. Uvular

The uvular sounds are produced by moving the root of the tongue near the uvula (Skandera & Burleigh, 2005). /q/ is uvular stop only in Urdu and Pashto. The sounds of /x/, /ɣ/ are uvular fricatives in Urdu (/ɣ/ with alternative symbol /g/) and uvular affricates in Sindhi.

Table 27: *Comparison of Uvular Consonants among Six Languages*

	Urdu Uvular	**Punjabi Uvular**	**Sindhi Uvular**	**Balochi Uvular**	**Pashto Uvular**	**Saraiki Uvular**
Stop	q				q	
Fricative	x					
Affricate			x ɣ			

2.10.10. Glottal

The glottal sounds are articulated in larynx when the air passed through the glottis (Skandera & Burleigh, 2005). All six languages, Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki have glottal fricative /h/ but only Sindhi language has voiced and voiceless /h/.

Table 28: *Comparison of Glottal Consonants among Six Languages*

	Urdu Glottal	**Punjabi Glottal**	**Sindhi Glottal**	**Balochi Glottal**	**Pashto Glottal**	**Saraiki Glottal**
Fricative	h	h	h h	h	h	h

2.11. Step II: Analysis of Phonetic Inventories on the basis of Manner of Articulation

According to Skandera and Burleigh, (2005) the manner of articulation refers the type of the closure by the speech organs. There are eight manners of articulation that use an egressive pulmonic air stream mechanism i.e. plosives, fricatives, affricates, nasals, laterals, trills, tap or flap and approximants. Along with manner of articulation continuant and non continuant difference is another terminological distinction for the analysis of the consonants. *Non-continuants* sounds are produced with the complete closure of the speech organs. Therefore plosives and affricates are non continuant sounds because are pronounced with the complete closure of oral or nasal tract. And all other sounds including nasal sounds are produced with the partial closure of the speech organs therefore are called *continuants*. Lateral sounds are also referred as *frictionless continuants* because there is no audible friction.

2.11.1. Stops/Plosives

The consonants which are produced by creating a short stricture (closure phase) by compressing the air in the mouth (hold phase) for a short time period than after holding (release phase) the air for a short time, the compressed air is released with the plosion (post release phase) are called stops or plosives. There are three types of plosives; bilabial plosives (/p/, /b/) palato alveolar plosives (/t/ and /d/) and velar plosives (/k/ and /g/) in English (Roach, 1998; Skandera & Burleigh, 2005).

/pʰ/ and /bʰ/ bilabial aspirated stops in Urdu, Sindhi and Saraiki languages, Punjabi has only voiceless bilabial stop /pʰ/. But the phonemic inventory of Urdu language has a large number of stop consonants than other languages such as /p/, /b/, /pʰ/, /bʰ/, /ʈ/, /ʈʰ/, /ɖ/, /ɖʱ/, etc. Balochi and Pashto have bilabial glide stop /w/. Only Sindhi and

Saraiki have bilabial ingressive stop /ɓ/. The sounds of /t̪/ and /d̪/are dental stops in Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki , /t̪ʰ / and /d̪ʰ / are aspirated dental stops present in Sindhi and Saraiki but Punjabi has only voiceless dental stop /t̪ʰ/. Balochi and Pashto have no aspirated labiodental stops. /ʈ/ and /ɖ/ are retroflex stop in Urdu, Punjabi, Balochi (with alternative symbols /ṭ/ and /ḍ/ respectively) Pashto and Saraiki. The sounds of /ʈʰ/ and /ɖʰ/ are aspirated retroflex stops in Urdu and Saraiki but Punjabi has only /ʈʰ/ voiceless aspirated retroflex stop. Retroflex implosive stop /ɗ/ is present only in Sindhi language. /k/ and /g/ are velar stops in all six languages; Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki . /kʰ/, /gʰ/ are aspirated velar stops in Urdu, Sindhi and Saraiki but Punjabi has only /kʰ / voiceless aspirated velar stop. Sindhi and Saraiki also have implosive velar stop /ɠ/. The uvular stop /q/ and glottal stop /ʔ/ are only present in Urdu and Pashto languages.

Table 29: *Comparison of Stops among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Bilabial Stop/ Plosives	p b pʰ bʰ	p pʰ b	p b pʰ bʰ	p b	p b	p b pʰ bʰ
Implosive			ɓ			ɓ
Dental Stop	t̪ d̪ t̪ʰ d̪ʰ	t̪ t̪ʰ d̪	t d tʰ dʰ	t d	t̪ d̪	t̪ d̪ t̪ʰ d̪ʰ
Retroflex Stop	ʈ ɖ ʈʰ ɖʰ	ʈ ʈʰ ɖ		ṭ ḍ	ʈ ɖ	ʈ ɖ ʈʰ ɖʰ
Velar Stop	k g kʰ gʰ	k kʰ g	k g kʰ gʰ	k g	k g	k g kʰ gʰ
Uvular Stop	q				q	
Glottal Stop	ʔ				ʔ	

2.11.2. Fricatives

The fricatives are produced with the release of air by a narrow passage with hissing sound and are also called *continuants* because one can continue to produce a fricative consonant without any interruption as long as one can sustain breath (Roach, 1998, pp. 47). According to Skandera and Burleigh (2005), in English, there are nine fricatives: /f/, /v/, /θ /, /ð/, /s/, /z/, /ʃ/, /ʒ/ and /h/. These fricative sounds are subdivided into two categories i.e. slit fricatives and groove fricatives depending on the width and depth of the tongue respectively. In English /f/, /v/, /θ /, /ð/ and /h/ are called ***slit fricatives*** because they have no point of narrowing and /s/, /z/, /ʃ/, /ʒ/ are also called ***groove fricatives*** by forming a groove along the tongue. Groove fricatives are pronounced with more intensity than slit fricatives therefore also called ***sibilants***.

/f/ is articulated as labiodental fricative in Urdu, Punjabi and Saraiki but is bilabial fricative in Pashto and Balochi. /v/ is labiodental fricative in Saraiki language. /ʧ/ and /ʤ/ are post alveolar fricative in Sindhi. /ʧʰ/ and /ʤʰ/ are post alveolar aspirated fricative in Sindhi, /ʃ/ is palatal fricative in Urdu, Punjabi, Saraiki and prepalatal palatal fricative in Balochi (with alternative symbol /š/) but is post alveolar fricative in Pashto. /ʒ/ is voiced palatal fricative in Urdu and prepalatal palatal voiced fricative in Balochi (with alternative symbol /ž/). The sounds of /x/ and /ɣ/ are uvular fricatives in Urdu (/ɣ/ with alternative symbol /g/), velar fricatives in Punjabi, Balochi (/ɣ/ with alternative symbol /ġ/), Pashto and Saraiki. /s/ and /z/ are alveolar fricative in Urdu, Sindhi and Pashto but are dental fricative in Punjabi, Balochi and Saraiki languages. /h/ is glottal fricative in all six languages (Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki) but Sindhi language has voiced glottal fricative /h/.

Table 30: *Comparison of Fricatives among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Bilabial Fricative				(f)	(f)	
Labiodental Fricative	f v	f v				f v
Dental Fricative			f v	s z		s z
Alveolar Fricative	s z	s z	s z			
Post-alveolar Fricative			ʧ ʤ ʧʰ ʤʰ			
Glottal Fricative	h	h	h h	h	h	h
Uvular Fricative	x ɡ					
Fricative Velar	ɣ	x ɣ		(x) (ġ)	x ɣ	x ɣ
Palatal Fricative	ʃ ʒ	ʃ		š ž		ʃ

2.11.3. Affricates

Affricates have mixed properties i.e. mixture of plosives and fricatives (Roach, 1998, pp. 52). According to Skandera and Burleigh, (2005) affricates are also called ***homorganic*** sounds. There are only two affricates in English; /ʧ/ and /ʤ/. They can also be regarded as *'groove fricatives'*.

/ʈ/, /ɖ/, /ʈʱ/ and /ɖʱ/ are non aspirated and aspirated retroflex affricates respectively in Sindhi language. /ʧ/ and /ʤ/ are alveolar affricate in Saraiki and palato alveolar in Pashto. /ʧʱ/ and /ʤʱ/ are aspirated palato alveolar affricates in Saraiki. /j/ is palatal affricate in Sindhi. /x/ and /ɣ/ are uvular affricates in Sindhi. /ʃ/ is post alveolar affricate in Sindhi.

Table 31: *Comparison of Affricates among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Dental Affricate					c j	
Alveolar Affricates	ʧ ʤ ʧʰ ʤʰ		ʃ			ʧ ʤ ʧʰ ʤʰ
Palato Alveolar Affricate					cˇ jˇ	
Retroflex Affricates			ʈ ɖ ʈʰ ɖʰ ɖʃ			
Palatal Affricates		ʧ ʤ ʧʰ	j	č ǰ		
Uvular Affricate			x ɣ			

2.11.4. Nasals

Nasal consonants are produced when air passes through the nasal tract by lowering the soft palate and creating a stricture in the oral cavity and the stricture may occur at three places i.e. bilabial stricture produces /m/, alveolar /n/ and velar /ŋ/ (Roach, 1998; Skandera & Burleigh, 2005).

The phoneme /m/ is present in all the six languages i.e. Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. /n/ is dental nasal in Urdu, Punjabi, Balochi and Saraiki, /n/ is alveolar voiced nasal in Sindhi. Retroflex nasal /ɳ/ is present only in Punjabi, Sindhi, Pashto (with an alternative symbol /ɳ/) and Saraiki. /ɳ/ is voiced in Sindhi and Pashto languages only. Sindhi, Pashto and Saraiki also have velar nasal /ɲ/. Sindhi language has another voiced palatal nasal /ɲ /.

Table 32: *Comparison of Nasals among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Bilabial Nasal	m	m	m	m	m	m
Dental Nasal	n	n		n		n̪
Alveolar Nasal			n			
Retroflex Nasal		ɳ	ɳ		ɳ	ɳ
Palatal Nasal			ɲ			
Nasal Velar			ɲ		ɲ	ɲ

2.11.5. Laterals

The lateral consonant /l/ is produced when the center of the tongue and alveolar ridge are in close contact by stopping the air passage than air releases from the sides of the tongue (Roach, 1998, pp. 58; Skandera & Burleigh, 2005). /l/ is an alveo-lateral in Urdu and Saraiki. /l/ is dento lateral in Punjabi, dento alveo lateral in Balochi and voiced alveolar tap in Pashto. Retroflex lateral /ɭ/ is only present in Punjabi and Sindhi languages. Aspirated lateral /lʰ / is observed in Urdu only.

Table 33: *Comparison of Laterals among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Dental Lateral		l		l		
Alveolar Lateral	l lʰ				l	l
Retroflex Lateral		ɭ	l			

2.11.6. Post alveolar Approximants

Post alveolar approximants are produced when articulators come close to each other without touching and releasing air to produce vowel like sounds and these are also called semi vowels, glides or approximants (Roach, 1998; pp. 59 & 60). But the gap between the articulators is larger than the fricatives (Skandera & Burleigh, 2005).

/j/ is palatal glide in Urdu, Pashto, Balochi (with alternative symbol /y/), palatal approximant in Saraiki and velar glide in Punjabi. /ɽ/ is retroflex glide in Pashto. /v/ is labiodental glide in Urdu. /r/ is voiced glide in Pashto. Post alveolar Implosive /ʃ/ is present only in Sindhi and Saraiki languages.

Table 34: *Comparison of Approximants among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Bilabial Approximant		ʋ		w	w	
Labiodental Approximant	ʋ					
Alveolar approximant					r	
Retroflex approximant					ɽ	

2.11.7. Trills/Rolls

Trill sound is produced when one articulator vibrates against another by forming an intermediate closure of speech organs in the vocal tract (Skandera & Burleigh, 2005). /r/ is retroflex trill in Sindhi language.

Table 35: *Comparison of Trills Consonants among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Retroflex Trill			r			

2.11.8. Tap/Flaps

Tap/Flaps are produced with the single flap (touch) of the tongue to the alveolar ridge (Skandera & Burleigh, 2005). /r/ is flap in Urdu, Punjabi, Balochi and Saraiki. Retroflex tap /ɽ/ is present in Urdu, Punjabi, Balochi (with alternative symbol /ɽ̤/) and Saraiki language but the aspirated retroflex tap /ɽʱ/ is rarely present only in Urdu.

Table 36: *Comparison of Flap/Taps Consonants among Six Languages*

	Urdu	**Punjabi**	**Sindhi**	**Balochi**	**Pashto**	**Saraiki**
Alveolar Tap/Flap	ɾ					ɾ
Dental Flap		ɾ		ɾ		
Retroflex Tap/Flap	ɽ ɽʰ	ɽ		ɽ̤		ɽ

The literature based comparison of the consonantal inventories of six major languages show;

- There are nine redundant alphabetical symbols in all six phonemic inventories of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. Most of these phonemes are borrowed from Persian and Arabic languages e.g. /s/ س, ص ث, /z/ ز ذ ظ , /t/ ط, and /?/ ع.
- In all six languages, there is large number of stop consonants in phonetic inventories than any other consonant. But among all six languages, Urdu language has a large number of stops.
- Regardless of other languages, Punjabi has no voiced aspirated stops but has only voiceless aspirated stops like /ʈʰ/, /ʈ ʱ/, etc. But Balochi and Pashto phonetic inventories have no aspirated stops.
- Sindhi language has a number of affricates among all six languages.
- There is no glide or approximant sound in Sindhi language.
- Only Sindhi language has trill /r/ sound.
- All six languages have glottal stop /h/ but only Sindhi language has both voiced and voiceless glottal stop /h/.
- Only Sindhi and Saraiki have ingressive/implosive sounds like /ɓ/, /ɠ/ etc.
- Retroflex nasal /ɳ/ is present only in Punjabi, Sindhi, Pashto (with an alternative symbol /ɳ̣/) and Saraiki. Sindhi, Pashto and Saraiki also have velar nasal /ŋ/. Sindhi language has another voiced palatal nasal /ɲ/.

The literature review is a brief report on:

- The origin, emerging socio political threats, geographical locations and the number of speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki in Pakistan and all over the world.
- The phonetic inventories of all these six languages i.e. Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki.
- The articulatory classification and description of the consonants and vowels of these languages.
- The phonetic similarities and differences among the inventories of these languages.
- The phonemic importance of these languages.
- The identification of the sounds of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages.
- The confused, redundant or interchanged sounds have also been discussed.

CHAPTER 3

METHODOLOGY

This research is based on quantitative methodology for analyzing accent variations in Pakistani Urdu. The phonetic inventories of six languages have been analyzed for finding out the similarities and gaps of segmental features of Corner vowels.

3.1. Nature of the Study

The research is exploratory in its nature and it has been designed to investigate the accent variations of Urdu language in Pakistan.

3.2. Population

Male and female speakers of the six majorly spoken languages in Pakistan

3.3. Sampling

30 speakers from each of the six languages (Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki) have been selected randomly as a sample.

3.4. Instrumentation

3.4.1. Instrument

Data has been collected by online telephone or mobile calls. PRAAT software (a package for analysis of speech in phonetics) has been used for analyzing the acoustical spectral distances by measuring F1, F2 differences. By using the recordings, perceptual analysis (accent identification) has been conducted based on vowels for comprehending different accents.

3.5. Corpus

The list of 139 districts' names of Pakistan has been selected as a corpus or vocabulary (list is also attached). The reason behind the selection of the list is; the district names are language independent and equally intelligible among the speakers of all six languages.

3.6. Data

720 recordings (30 speakers x 6 languages x 4 corner vowels = 720 utterances) have been used for analyzing the vowels (which are available through the data). All these vowels arising from Urdu L2 gap, would find out pair wise and group wise comparison.

3.7. Data Analysis

The data analysis has been conducted on the acoustic phonetic analysis of the vocalic inventories.

Word List Used for Data Collection and Analysis

List of 139 district names in Pakistan used as Corpus/data for recordings

(For CISAMPA see Appendix C)

WORD		CISAMPA
1. ABBOTTABAD	ایبٹ آباد	A_EBATA_ABA_AD
2. ASTORE	استور	AST_DU_UR
3. ATTOCK	اٹک	ATAK
4. AWARAN	آواران	A_AVA_ARA_AN
5. BADIN	بدین	BAD_DI_IN
6. BAGH	باغ	BA_AG_G
7. BAHAWALNAGAR	بہاولنگر	BAHA_AVALNAGAR
8. BAHAWALPUR	بہاولپور	BAHA_AVALPU_UR
9. BAJAUR	باجور	BA_AD_ZO_OR_R
10. BANNU	بنوں	BANNU_U_N
11. BARKHAN	برکھان	BARK_HA_AN
12. BATTAGRAM	بٹگرام	BATGARA_AM
13. BHAKKAR	بھکر	B_HAKKAR
14. BHIMBER	بھمبھر	B_HIMBAR
15. BOLAN	بولان	BO_OLA_AN
16. BUNER	بونیر	BU_UNI_IR
17. CHAGAI	چاغی	T_SA_AG_GI_I
18. CHAKWAL	چکوال	T_SAKVA_AL
19. CHARSADA	چارسدہ	T_SA_ARSAD_DD_DA_A

WORD	**CISAMPA**
20. CHINIOT چنیوٹ	T_SANJO_OT
21. CHITRAL چترال	T_SIT_DRA_AL
22. DADU دادو	D_DA_AD_DU_U
23. DERA_BUGTI ڈیرہ بگٹی	DA_YRA_ABUGTI_I
24. DERA_GHAZI_KHAN ڈیرہ غازی خان	DA_YRA_AG_GA_AZI_IXA_AN
25. DERA_ISMAIL_KHAN ڈیرہ اسمعیل خان	DA_YRA_AISMA_AI_ILXA_AN
26. DIAMER دیا مر	DIJA_AMAR
27. FAISALABAD فیصل آباد	FA_ESALA_ABA_AD_D
28. GHANCHE غانچے	G_GA_ANT_SA_Y
29. GHIZER غزر	G_GAZAR
30. GHOTKI گھوٹکی	G_HO_OTKI_I
31. GILGIT گلگت	GILGIT_D
32. GUJRANWALA گجرنوالہ	GUD_ZRA_A_NVALA_A
33. GUJRAT گجرات	GUD_ZRA_AT_D
34. GWADAR گوادر	GAVA_AD_DAR
35. HAFIZABAD خافظ اباد	HA_AFIZA_ABA_AD_D
36. HANGU ہنگو	HAN_GGU_U
37. HARIPUR ہری پور	HARI_IPU_UR
38. HARNAI ہرنائی	HARNA_AI_I
39. HATTIAN ہٹیاں	HATTIA_A_N

WORD	CISAMPA
40. HAVELI حویلی	HAVA_YLI_I
41. HUNZA_NAGAR ہنزہ نگر	HUNZA_ANAGAR
42. HYDERABAD حیدر آباد	HA_ED_DARA_ABA_AD_D
43. ISLAMABAD اسلام آباد	ISLA_AMA_ABA_AD_D
44. JACOBABAD جیکب آباد	D_ZA_EKABA_ABA_AD_D
45. JAFARABAD جعفر آباد	D_ZA_AFARA_ABA_AD_D
46. JAMSHORO جامشورو	D_ZA_AMS_HO_ORO_O
47. JANUBI_WAZIRISTAN جنوبی وزیرستان	D_ZANU_UBI_IVAZI_IRIST_DA_AN
48. JHAL_MAGSI جھل مگسی	D_Z_HALMAGSI_I
49. JHANG جھنگ	D_Z_HAN_GG
50. JHELUM جہلم	D_ZA_Y_HHLAM
51. KALAT قلات	QALA_AT_D
52. KAMBAR_SHAHDADKOT قمبر شاہ داد کوٹ	QAMBARS_HAHD_DA_AD_DKO_OT
53. KARACHI کراچی	KARA_AT_SI_I
54. KARAK کرک	KARAK
55. KASHMORE کاشمور	KAS_HMO_OR
56. KASUR قصور	QASU_UR
57. KECH کیچ	KA_YT_S
58. KHAIRPUR خیر پور	XA_ERPU_UR
59. KHANEWAL خانیوال	XA_ANA_YVA_AL

WORD		CISAMPA
60. KHARAN	خاران	XA_ARA_AN
61. KHUSHAB	خوشاب	XUS_HA_AB
62. KHUZDAR	خنزه	XUZD_DA_AR
63. KHYBER	خیبر	XA_EBAR
64. KILLA_ABDULLAH	قلعہ عبداللہ	QALA_AABDULLA_A_H
65. KILLA_SAIFULLAH	قلعہ سیف اللہ	QALA_ASA_EFULLA_A_H
66. KOHAT	کوہاٹ	KO_OHA_AT
67. KOHISTAN	کوہستان	KO_OHIST_DA_AN
68. KOHLU	کوہلو	KO_OHLU_U
69. KOTLI	کوٹلی	KO_OTLI_I
70. KURRAM	قرم	QURRAM
71. LAHORE	لاہور	LA_AHO_OR
72. LAKKI_MARWAT	لکی مروت	LAKKI_IMARVAT_D
73. LARKANA	لاڑکانہ	LA_AR_RKA_ANA_A
74. LASBELA	لسبیلہ	LASBA_YLA_A
75. LAYYAH	لیہ	LAJJA_AH
76. LODHRAN	لودراں	LO_OD_D_HRA_A_N
77. LORALAI	لورالائی	LO_ORA_ALA_AI_I
78. LOWER_DIR	لوئر دیر	LO_OIRD_DI_IR
79. MALAKAND	مالاکنڈ	MA_ALA_AKAND
80. MANDI_BAHAUDIN	منڈی بہاوٗدی	MANDI_IBAHA_AUD_DD_DI_IN
81. MANSEHRA	مانسہرہ	MA_ANSA_Y_HHRA_A

WORD	CISAMPA
82. MARDAN مردان	MARDA_AN
83. MASTUNG مستونگ	MAST_DU_U_NN_GG
84. MATIARI مٹیاری	MATIA_ARI_I
85. MIANWALI میانوالی	MIA_A_NVA_ALI_I
86. MIRPUR میر پور	MI_IRPU_UR
87. MIRPURKHAS میرپور خاص	MI_IRPU_URXA_AS
88. MOHMAND موہمند	MO_HMAND_D
89. MULTAN ملتان	MULT_DA_AN
90. MUSAKHEL موسی خیل	MU_USA_AXA_YL
91. MUZAFFARABAD مظفر آباد	MUZAFFARA_ABA_AD_D
92. MUZAFFARGARH مظفر گڑھ	MUZAFFARGAR_R_H
93. NANKANA_SAHIB ننکانہ صاحب	NANKA_ANA_ASA_AHIB
94. NAROWAL نارووال	NA_ARO_OVA_AL
95. NASIRABAD ناصر آباد	NASI_IRA_ABA_AD_D
96. NAUSHAHRO_FIROZE نوشیروفیروز	NO_OS_HA_E_HHRO_OFI_IRO_OZ
97. NEELUM نیلم	NI_ILAM
98. NOWSHERA نوشیرہ	NO_OS_HA_E_HHRA_A
99. NUSHKI نشکی	NO_OS_HKI_I
100. OKARA اکاڑہ	O_OKA_AR_RA_A
101. ORAKZAI اورکزئی	ORAKZAI_I
102. PAKPATTAN پاکپتن	PA_AKPAT_DAN

WORD			CISAMPA
103.	PANJGUR	پنج گور	PAND_ZGU_UR
104.	PESHAWAR	پشاور	PIS_HA_AVAR
105.	PISHIN	پشین	PAS_HI_IN
106.	POONCH	پونچھ	PU_U_NT_S_H
107.	QUETTA	کوئٹہ	KO_OA_Y_HTA_A
108.	RAHIM_YAR_KHAN	رحیم یار خان	RAHI_IMYA_ARXA_AN
109.	RAJANPUR	راجن پور	RA_AD_ZANPU_UR
110.	RAWALPINDI	راولپنڈی	RA_AVALPINDI_I
111.	SAHIWAL	ساہیوال	SA_AHI_IVA_AL
112.	SANGHAR	سانگھڑ	SA_A_NG_HAR_R
113.	SARGODHA	سرگودھا	SARGO_OD_D_HA_A
114.	SHAHEEDBENAZIRABAD	شہید بینظیر آباد	S_HAHI_ID_DBA_YNAZI_IRA_ABA_AD
115.	SHANGLA	شانگلہ	S_HA_A_NN_GGLA_A
116.	SHEIKHUPURA	شیخوپورہ	S_HA_YXU_UPU_URA_A
117.	SHERANI	شیرانی	S_HA_YRA_ANI_I
118.	SHIKARPUR	شکار پور	S_HIKA_ARPU_UR
119.	SHUMALI_WAZIRISTAN	شمالی وزیرستان	S_HUMA_ALI VAZI_IRIST_DA_AN
120.	SIALKOT	سیالکوٹ	SA_Y_HA_ALKO_OT
121.	SIBI	سبی	SIBBI_I

WORD			CISAMPA
122.	SKARDU	سکردو	SAKARD_DU_U
123.	SUDHNUTTI	سدھ نوتی	SUD_D_HNU_UT_DI_I
124.	SUKKUR	سکھر	SAK_HK_HAR
125.	SWABI	سوابی	SAVA_ABI_I
126.	SWAT	سوات	SAVA_AT_D
127.	TANDO_ALLAHYAR	ٹنڈو الله یار	TANDU_UALLA_HJA_AR
128.	TANDO_MUHAMMADKHAN	ٹنڈو محمد خان	TANDU_UMO_O_HHAMMAD_DXA_AN
129.	TANK	تنک	T_DA_A_NN_GK
130.	THARPARKAR	تھرپارکر	T_HARPA_ARKAR
131.	THATTA	ٹھٹھہ	T_HAT_HT_HA_A
132.	TOBA_TEK_SINGH	ٹوبہ ٹیک سنگ	TO_OBA_ATA_YKSIN_GG_H
133.	TOR_GHAR	تورگر	T_DO_ORG_HAR
134.	UMERKOT	عمر کوٹ	UMARKO_OT
135.	UPPER_DIR	اپر دیر	APPARD_DI_IR
136.	VEHARI	ویہاڑی	VA_Y_HHA_AR_RI_I
137.	WASHUK	واشوک	VAS_HO_OK
138.	ZHOB	ژوب	Z_ZOB
139.	ZIARAT	زیارت	ZA_Y_HA_ARAT_D

CHAPTER 4

DATA ANALYSIS & RESULTS

The data analysis of vocalic inventories has been conducted in PRAAT. The acoustic phonetic analysis of corner vowels (uttered by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages) has been conducted for finding out the accent variation in Pakistani Urdu language.

4. Acoustic Analysis of Vowels

Phonetically, the vowels sounds are produced without any obstruction of the air in oral or nasal tract and phonologically vowels occupy the central position in the syllable (Skandera & Burleigh, 2005). According to Ladefoged, vowels are produced without causing friction in the oral tract. In comparison with the consonant sounds, vowel sounds are difficult to study and required more attention than consonants (as cited in Kerio, 2010; Kent, 2002; Moore, 2003).

In the present research, only the characteristics of vowels have been investigated to identify the accents of six major languages; Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki in Pakistan. The 139 district names of Pakistan have been used as a speech corpus (for details see Chapter no. 3). This corpus covered all the vowels in the six languages. The corpus was recorded by male and female native speakers with ages varying from 18 to 40 years old. The speakers' distribution in each accent is listed in table no. 38.

Table 37: *Speaker Distribution in each Accent*

Accents	**I_I**		**A_E**		**A_A**		**U_U**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Urdu	22	8	11	9	17	13	14	16
Punjabi	17	13	16	4	17	13	22	8
Sindhi	12	18	10	8	12	18	7	23
Balochi	15	5	8	2	17	13	21	2
Pashto	27	3	20	7	25	5	27	3
Saraiki	17	13	16	11	17	13	18	12

The speech corpus has been recorded over telephone or mobile channel at 8000Hz from the students of the different universities in Pakistan with random background noise. For the experiment, the four corner vowels[41] /i:/, /æ:/, /a:/ and /u:/ have been used as fixed reference points for the description of other vowels in the vocalic inventories. The vowel data used for this study is given in table 39.

Table 38: *Accent Variation among Six Languages*

Accents	**I_I (i:)**	**A_E (æ:)**	**A_A (a:)**	**U_U (u:)**
Urdu	30	20	30	30
Punjabi	30	20	30	30
Sindhi	30	18	30	30
Balochi	20	10	30	23
Pashto	30	27	30	30
Saraiki	30	27	30	30

[41] According to Oxford Urdu-English Dictionary 2013, duration mark /:/ has been used to differentiate between a long vowel and a short vowel

Formants are sensitive features of a (Chinese) language (Stantic & Jo, 2012). So in our work, we use formants F1, F2 to classify the six major accents of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages in Pakistan. Formants have been manually calculated by using PRAAT software. After normalizing the data, the mean vector for each vowel accent have been computed. Each mean value represents an individual language accent of a specific vowel. Then the distance between two accents of a vowel has been computed by using mean vectors of these accents. And for confirming the results, one way ANOVA statistical analysis has been done by using SPSS software (for details see APPENDIX E).

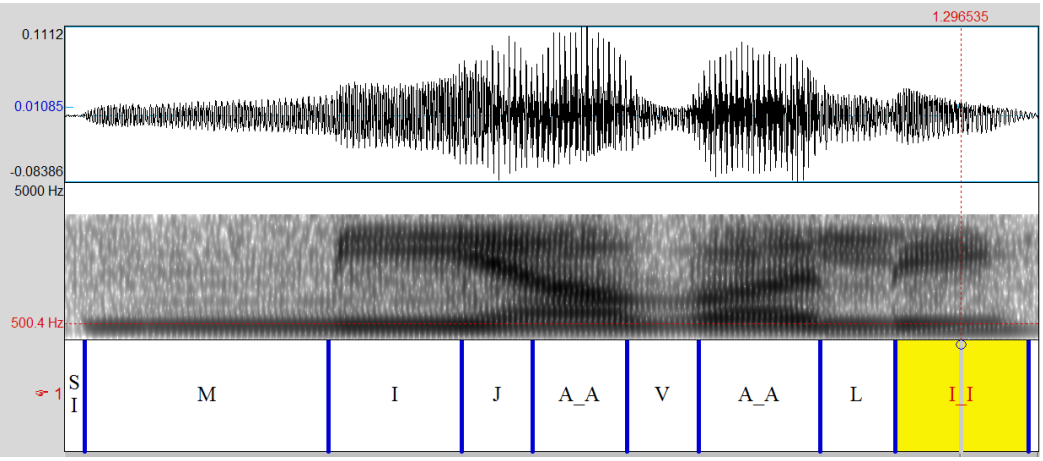
The accent variation among six major languages; Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki has been measured by taking frequencies. In the first step the wave files have been trimmed by excluding background noise from the required voiced signals. In the second step the spectrogram of vowel (districts' names) utterances have been computed and F1, F2 frequencies of the vowels have been marked manually by viewing and measuring spectral slices. In an utterance, formants are usually present in all voiced phonemes such as vowels, approximants and nasals but formants are stable and strong in vowels over the time (as cited in Keerio, 2010; Moore 2003). Urdu formant frequencies can be defined within these ranges; F1 ranges between 200-700 Hz and F2 ranges between 1100-2400 Hz for an adult speaker.

The used symbols for vowels are similar in different languages but their properties are not identical. So, the reference point is useful despite of the similar transcription. Therefore, the phonetics system allowed the languages to describe the vowels with relation to the acoustic properties despite the height of the tongue

(McCombs, 2006) but vowel formants are more dependent on lips movement (Hillenbrand, Getty, Clark, & Michael, 1995). In this study, the acoustic properties of corner vowels in Urdu have been analyzed by keeping in mind the above mentioned properties.

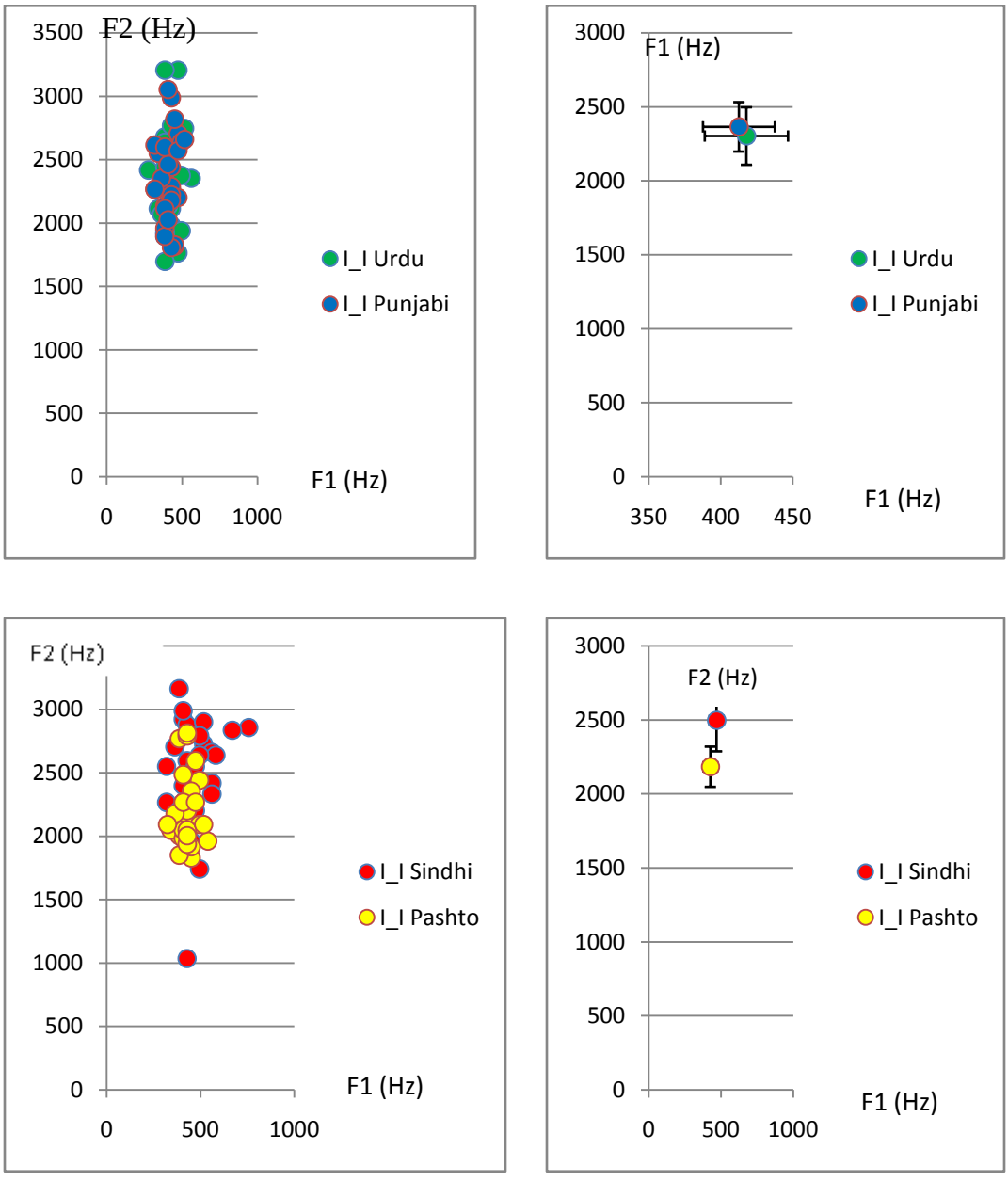
4.1. Acoustic Analysis of I_I (i:) Vowels

/i:/ vowel is produced when the tongue pushes forward by minimizing the gap between the tongue body and the palate but lips remain neutral for the articulation (Skandera & Burleigh, 2005). In Urdu language, /i:/ vowel has lower F1 and higher F2 values (200-2400 Hz) than /a:/ vowel.



The recorded data was processed in PRAAT software. The data is based on the F1, F2 values of /i:/ vowels. For the understanding of the data variation cluster graphs have been made which show the similar as well as contrastive behavior of different vowels uttered by the speakers of six languages. These graphical representations show that frequencies of vowels vary from each other but some pair wise comparisons show more contrastive behavior than the other. Few of them are elaborated here;

Pair wise Comparison of I_I Vowels Uttered by Different Language Speakers



The pair wise comparison shows rate of variation in the formant frequencies (F1, F2) among the utterances from the speakers of the different languages. The pair wise comparison of I_I vowel uttered by Urdu and Punjabi speakers shows less variation than the I_I vowel uttered by Sindhi and Pashto speakers.

Figure 25: shows the group wise comparison of the mean and standard deviation of (the long vowel /i:/) I_I uttered by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages. This group wise comparison shows that the formant frequency values of I_I vowel uttered by the Balochi speakers are drastically different from the formant frequencies of I_I vowel uttered by the speakers of Sindhi, Punjabi, Urdu and Saraiki but also shows difference from the formant frequencies of I_I vowels uttered by Pashto speakers.

Group wise Comparison of I_I Vowels

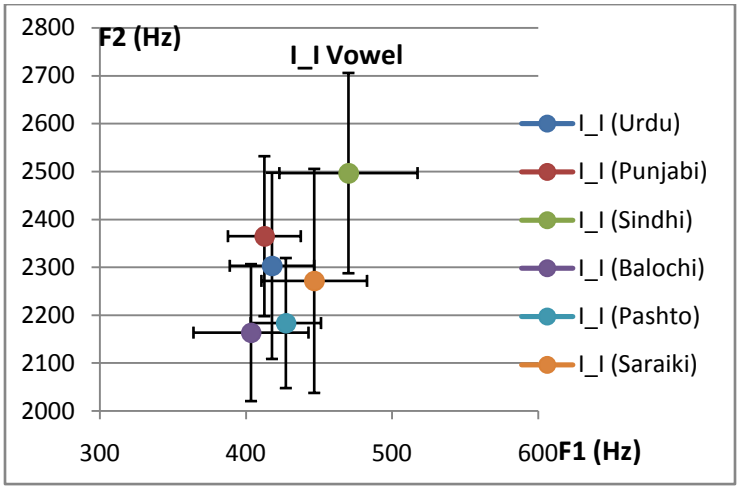


Figure 25: Mean & St. Deviation of /i:/ Vowel

ANOVA test has been run for the verification of the results.

Statistical Analysis of I_I (i:) Vowel

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
I_I.F1	Between Groups	84432.180	5	16886.436	3.620	.004
	Within Groups	765060.267	164	4665.002		
	Total	849492.447	169			
I_I.F2	Between Groups	2086300.039	5	417260.008	3.011	.013
	Within Groups	22730305.967	164	138599.427		
	Total	24816606.006	169			

Multiple Comparisons of I_I (i:) Vowel

Dependent Variable	(I) Group	(J) Group	Sig.	Dependent Variable	(I) Group	(J) Group	Sig.
I_I.F1	Urdu	Punjabi	1.000	I_I.F2	Urdu	Punjabi	.987
		Sindhi	.039			Sindhi	.339
		Balochi	.978			Balochi	.786
		Pashto	.994			Pashto	.815
		Saraiki	.573			Saraiki	.999
	Punjabi	Urdu	1.000		Punjabi	Urdu	.987
		Sindhi	.017			Sindhi	.745
		Balochi	.997			Balochi	.422
		Pashto	.961			Pashto	.414
		Saraiki	.384			Saraiki	.926
	Sindhi	Urdu	.039		Sindhi	Urdu	.339
		Punjabi	.017			Punjabi	.745
		Balochi	.011			Balochi	.027
		Pashto	.152			Pashto	.017
		Saraiki	.769			Saraiki	.184
	Balochi	Urdu	.978		Balochi	Urdu	.786
		Punjabi	.997			Punjabi	.422
		Sindhi	.011			Sindhi	.027
		Pashto	.831			Pashto	1.000
		Saraiki	.245			Saraiki	.916
	Pashto	Urdu	.994		Pashto	Urdu	.815
		Punjabi	.961			Punjabi	.414
		Sindhi	.152			Sindhi	.017
		Balochi	.831			Balochi	1.000
		Saraiki	.880			Saraiki	.942
	Saraiki	Urdu	.573		Saraiki	Urdu	.999
		Punjabi	.384			Punjabi	.926
		Sindhi	.769			Sindhi	.184
		Balochi	.245			Balochi	.916
		Pashto	.880			Pashto	.942

Table 39: *Difference among /i:/ Vowels*

Significant Values		
Accents	**F1**	**F2**
Urdu- Sindhi	0.039	---
Punjabi- Sindhi	0.017	---
Sindhi-Balochi	0.011	.027
Sindhi-Pashto	---	.017

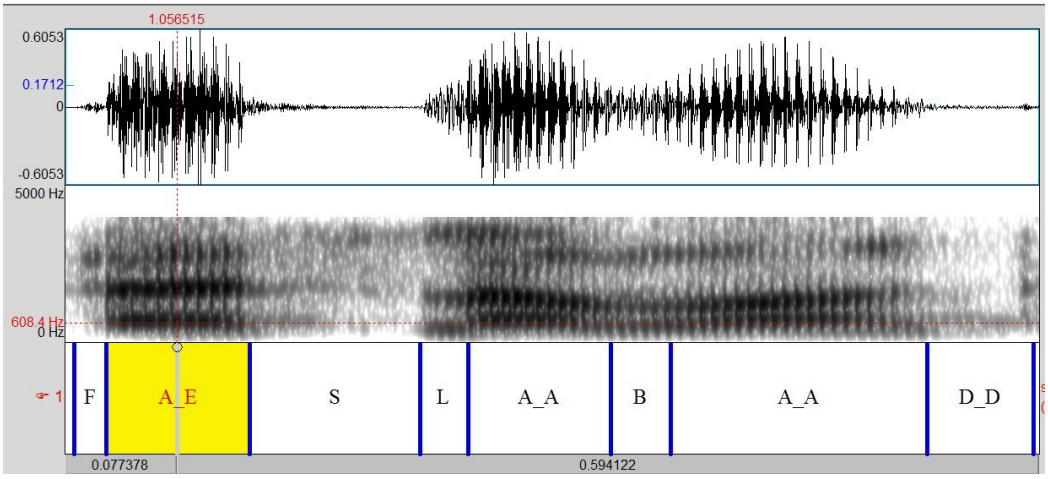
One way ANOVA test was run to determine, if there was any significant difference in I_I vowel accent (F1, F2) among the six groups of languages. There were no outliers and the data was normally distributed for each group and assessed by the Tukey post hoc test [$p < .05$]. There was statistically significant difference among the groups as determined by one way ANOVA [$F (5, 164) = 3.62$, $P = .004$]. A Tukey *post-hoc* test revealed that F1 of I_I (i:) vowel in Urdu is significantly different from Sindhi I_I [$P = .039$], Punjabi I_I vowel is significantly different from Sindhi [$P = .017$] I_I vowel. Sindhi I_I vowel is statistically significantly different from Balochi [$P = .011$] I_I vowel. There was statistically significant difference among the six language groups as determined by one way ANOVA [$F (5, 164) = 3.011$, $P = .013$]. A Tukey *post-hoc* test revealed that F2 of I_I (i:) vowel. Sindhi I_I vowel is significantly different from Balochi [$P = .027$] and Pashto [$P = .017$] I_I vowels (for details see APPENDIX E).

The analysis shows that F1, F2 values of I_I vowels among all these six accents are different from each other i.e. I_I vowel uttered by Urdu and Punjabi speakers are significantly different from I_I vowel uttered by Sindhi speakers. I_I vowels uttered by

Sindhi speakers are significantly different from I_I vowels uttered by Balochi and Pashto speakers.

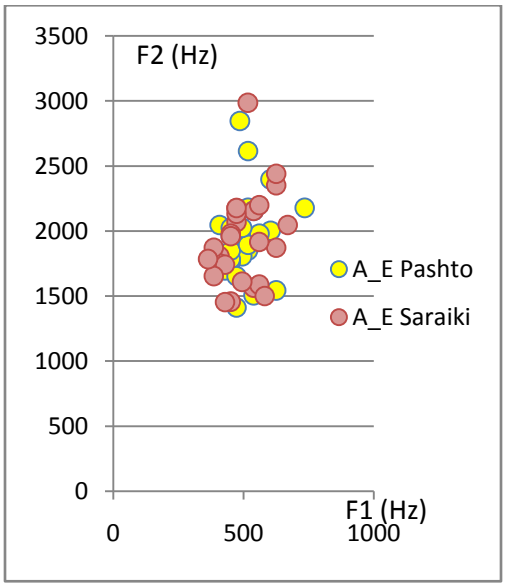
4.2. Acoustic Analysis of A_E (æ:) Vowels

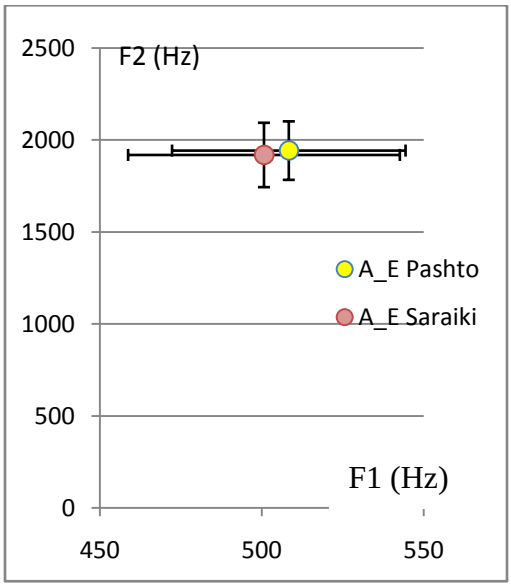
/æ:/ vowel is produced when the tongue moves backward by increasing the gap between lower and upper jaws but lips are in more spread posture than /i:/ vowel (Skandera & Burleigh, 2005). In Urdu, /æ:/ vowel has slightly lower F1 value and higher F2 value (700-1700 Hz) than /a:/ vowel.

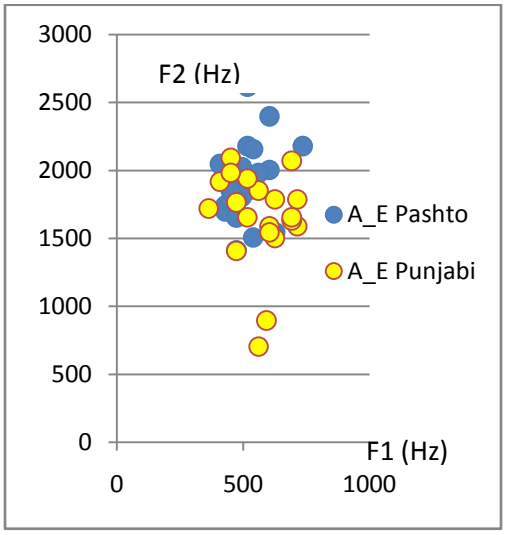


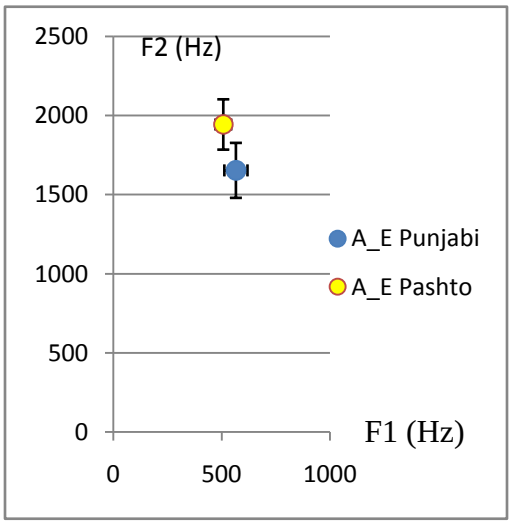
The recorded data was processed in PRAAT software. The data is based on F1, F2 values of /æ:/ vowels. For the understanding of the data variation the cluster graphs have been made which show the similar as well as contrastive behavior of different vowels uttered by the speakers of six languages. These graphical representations show that frequencies of vowels vary from each other but some pair wise comparisons show more contrastive behavior than the other. Few of them are elaborated here;

Pair wise Comparison of A_E Vowels Uttered by Different Language Speakers









The pair wise comparison shows rate of variation in the formant frequencies (F1, F2) among the utterances from the speakers of the different languages. The pair wise comparison of A_E vowel uttered by Pashto and Saraiki speakers shows less variation than A_E vowel uttered by Punjabi and Pashto speakers.

Figure 26: shows the group wise comparison of the mean and standard deviation of (the long vowel /æ:/) A_E vowel uttered by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages. This group wise comparison shows that the formant frequency values of A_E vowel uttered by the Balochi speakers are equally different from the formant frequencies of A_E vowel uttered by the speakers of Urdu, Punjabi, Sindhi, Saraiki and Pashto. The formant frequencies of A_E vowel uttered by Punjabi speakers are drastically different from Sindhi, Pashto, Saraiki and Balochi but less drastically variant from the utterances of Urdu speakers.

Group wise Comparison of A_A Vowels

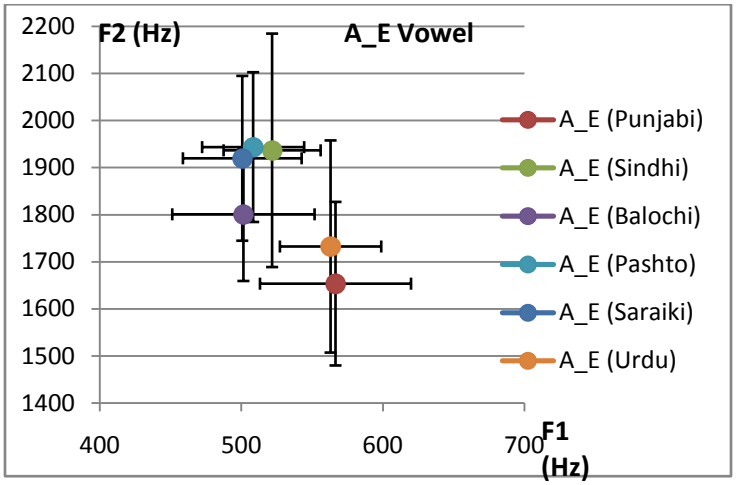


Figure 26: Mean & St. Deviation of /æ:/ Vowels

ANOVA test has been run for the verification of the results.

Statistical Analysis of A_E (æ:) Vowel

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
A_E.F1	Between Groups	92435.684	5	18487.137	2.669	.025
	Within Groups	803392.427	116	6925.797		
	Total	895828.112	121			
A_E.F2	Between Groups	1569726.338	5	313945.268	2.653	.026
	Within Groups	13726734.843	116	118333.921		
	Total	15296461.180	121			

Multiple Comparisons of A_E (æ:) Vowel

Dependent Variable	(I) Group	(J) Group	Sig.	Dependent Variable	(I) Group	(J) Group	Sig.
A_E.F1	Urdu	Punjabi	1.000	A_E.F2	Urdu	Punjabi	.978
		Sindhi	.647			Sindhi	.453
		Balochi	.401			Balochi	.996
		Pashto	.233			Pashto	.305
		Saraiki	.121			Saraiki	.441
	Punjabi	Urdu	1.000		Punjabi	Urdu	.978
		Sindhi	.562			Sindhi	.123
		Balochi	.338			Balochi	.878
		Pashto	.175			Pashto	.056
		Saraiki	.086			Saraiki	.100
	Sindhi	Urdu	.647		Sindhi	Urdu	.453
		Punjabi	.562			Punjabi	.123
		Balochi	.990			Balochi	.917
		Pashto	.995			Pashto	1.000
		Saraiki	.961			Saraiki	1.000
	Balochi	Urdu	.401		Balochi	Urdu	.996
		Punjabi	.338			Punjabi	.878
		Sindhi	.990			Sindhi	.917
		Pashto	1.000			Pashto	.872
		Saraiki	1.000			Saraiki	.937
	Pashto	Urdu	.233		Pashto	Urdu	.305
		Punjabi	.175			Punjabi	.056
		Sindhi	.995			Sindhi	1.000
		Balochi	1.000			Balochi	.872
		Saraiki	.999			Saraiki	1.000
	Saraiki	Urdu	.121		Saraiki	Urdu	.441
		Punjabi	.086			Punjabi	.100
		Sindhi	.961			Sindhi	1.000
		Balochi	1.000			Balochi	.937
		Pashto	.999			Pashto	1.000

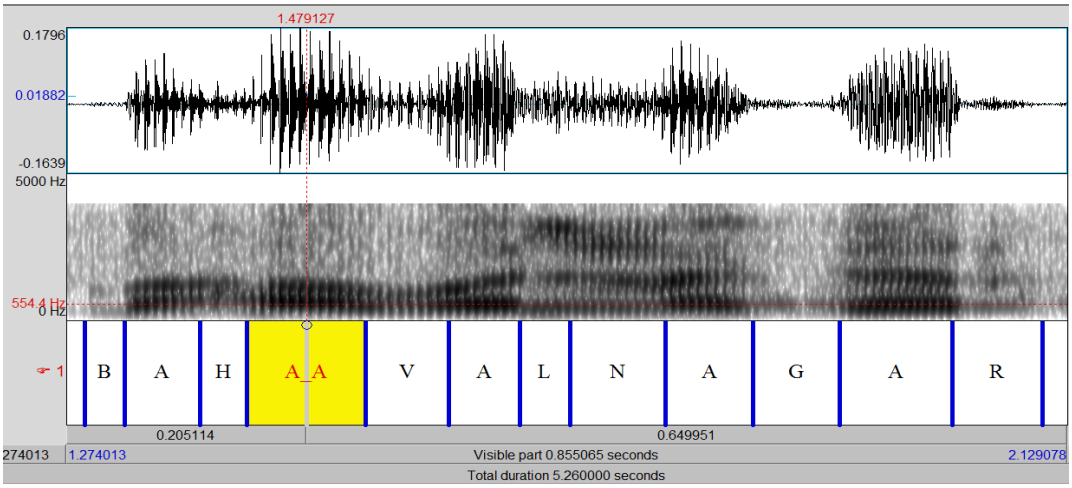
Table 40: *Difference among /æ:/ vowels*

Accents	**F1**
Punjabi-Pashto	0.056

One way ANOVA test was run to determine, if there was any significant difference in A_E vowel accent (F1, F2) among the six groups of languages. There were no outliers and the data was normally distributed for each group and assessed by the Tukey *post-hoc* test [p < .05]. There was statistically significant difference among groups as determined by one way ANOVA [F (5, 116) = 2.669, *P* = .025]. A Tukey *post-hoc* test revealed that F1 of A_E (æ:) vowel in Punjabi is significantly different from Pashto [*P* = .056] A_E vowels (for details see APPENDIX E). The result shows that there is no significant difference among other formant frequencies of the vowels except Punjabi and Pashto. So, A_E vowel in Urdu and Punjabi accents are not significantly different from each other. But A_E vowels uttered by the Punjabi speakers are significantly different from the A_E vowel uttered by Pashto speakers.

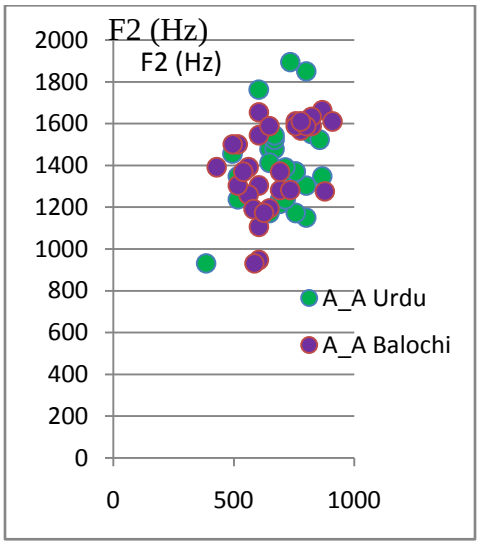
4.3. Acoustic Analysis of A_A (a:) Vowel

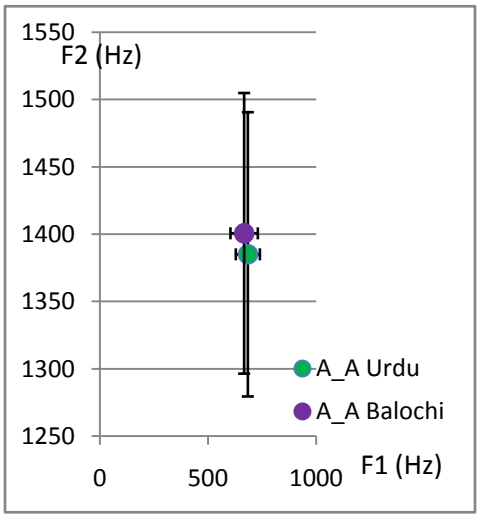
/a:/ vowel is produced when the tongue moves backward by maximizing the gap between lower and upper jaws but lips remain wide open for articulation (Skandera & Burleigh, 2005). In Urdu language, /a:/ vowel has upper F1 and lower F2 values (700-1100 Hz) than /i:/ vowel.

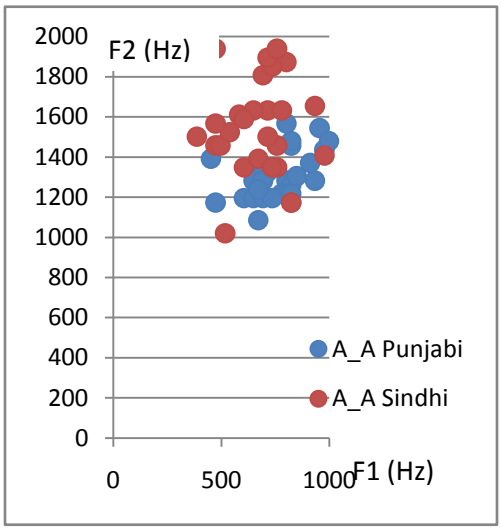


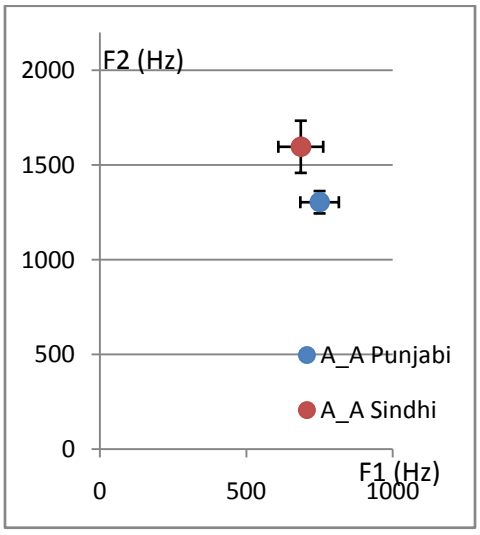
The recorded data was processed in PRAAT software. The data is based on the F1, F2 values of /a:/ vowels. For the understanding of the data variation cluster graphs have been made which show the similar as well as contrastive behavior of different vowels uttered by the speakers of six languages. These graphical representations show that frequencies of vowels vary from each other but some pair wise comparisons show more contrastive behavior than the other. Few of them are elaborated here;

Pair wise Comparison of A_A Vowels Uttered by Different Language Speakers









The pair wise comparison shows rate of variation of the formant frequencies (F1, F2) among the utterances from the speakers of the different languages. The pair wise comparison of A_A vowel uttered by Urdu and Balochi speakers shows less variation than the A_A vowel uttered by Punjabi and Sindhi speakers.

Figure 27: shows the group wise comparison of the mean and standard deviation of (the long vowel /a:/) A_A vowels uttered by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages. This group wise comparison shows that the formant frequency values of A_A vowel uttered by Punjabi speakers are drastically different from the formant frequencies of A_A vowels uttered by the speakers of Sindhi language. A_A vowel uttered by Punjabi speakers shows deviation from A_A vowels uttered by the speakers of Urdu, Balochi, Pashto and Saraiki languages. A_A vowels uttered by the speakers of Sindhi language are totally different from A_A vowel uttered by the speakers of Urdu, Punjabi, Balochi, Pashto and Saraiki languages.

Group wise Comparison of A_A Vowels

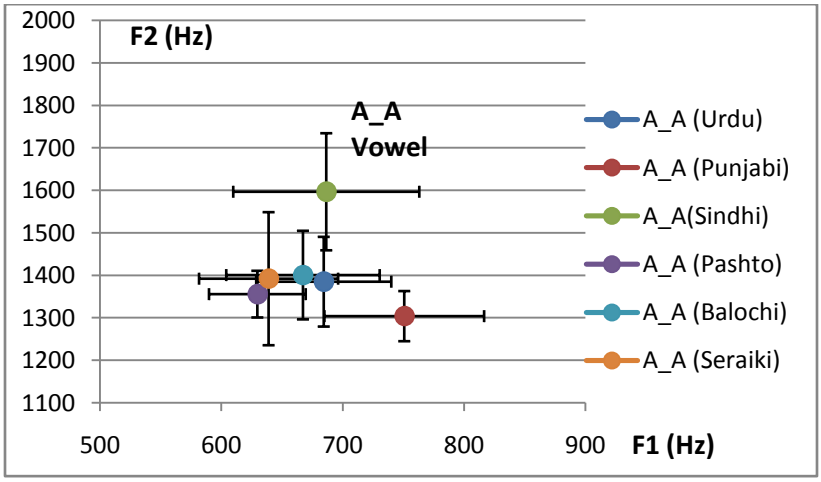


Figure 27: Mean & St. Deviation of /a:/ Vowels

ANOVA test has been run for the verification of the results.

Statistical Analysis of A_A (a:) Vowel

ANOVA						
		Sum of Squares	Df	Mean Square	F	Sig.
A_AF1	Between Groups	280854.311	5	56170.862	3.799	.003
	Within Groups	2572575.267	174	14784.915		
	Total	2853429.578	179			
A_AF2	Between Groups	1498418.517	5	299683.703	6.244	.000
	Within Groups	8350693.233	174	47992.490		
	Total	9849111.750	179			

Multiple Comparisons of A_A (a:) Vowel

Dependent Variable	(I) Group	(J) Group	Sig.	Dependent Variable	(I) Group	(J) Group	Sig.
A_AF1	Urdu	Punjabi	.286	A_AF2	Urdu	Punjabi	.707
		Sindhi	1.000			Sindhi	.003
		Balochi	.994			Balochi	1.000
		Pashto	.505			Pashto	.995
		Saraiki	.697			Saraiki	1.000
	Punjabi	Urdu	.286		Punjabi	Urdu	.707
		Sindhi	.320			Sindhi	.000
		Balochi	.089			Balochi	.529
		Pashto	.002			Pashto	.942
		Saraiki	.006			Saraiki	.628
	Sindhi	Urdu	1.000		Sindhi	Urdu	.003
		Punjabi	.320			Punjabi	.000
		Balochi	.990			Balochi	.009
		Pashto	.463			Pashto	.000
		Saraiki	.656			Saraiki	.005
	Balochi	Urdu	.994		Balochi	Urdu	1.000
		Punjabi	.089			Punjabi	.529
		Sindhi	.990			Sindhi	.009
		Pashto	.838			Pashto	.968
		Saraiki	.945			Saraiki	1.000
	Pashto	Urdu	.505		Pashto	Urdu	.995
		Punjabi	.002			Punjabi	.942
		Sindhi	.463			Sindhi	.000
		Balochi	.838			Balochi	.968
		Saraiki	1.000			Saraiki	.988
	Saraiki	Urdu	.697		Saraiki	Urdu	1.000
		Punjabi	.006			Punjabi	.628
		Sindhi	.656			Sindhi	.005
		Balochi	.945			Balochi	1.000
		Pashto	1.000			Pashto	.988

One way ANOVA test was run to determine, if there was any significant difference in A_A vowel accent (F1, F2) among the six groups of language. There were no outliers and the data was normally distributed for each group and assessed by the Tukey *post hoc* test [$P < .05$]. There was statistically significant difference among groups as determined by one way ANOVA [$F (5, 174) = 3.799$, $P = .003$]. A Tukey *post-hoc* test revealed that F1 of A_A (a:) vowel in Punjabi is significantly different from Pashto accent [$P = .002$]. There was statistically significant difference among groups as determined by one way ANOVA [$F (5, 174) = 6.244$, $P = .000$]. A Tukey *post-hoc* test revealed that F2 of A_A (a:) vowel uttered by the speakers of Urdu is significantly different from the speakers of Sindhi language [$P = .003$], A_A vowels uttered by the speakers of Punjabi language are significantly different from the speakers of Sindhi language [$P = .000$], A_A vowels uttered by the speakers of Sindhi language are significantly different from the speakers of Pashto [$P = .000$] and Saraiki [$P = .005$]. For details see APPENDIX E

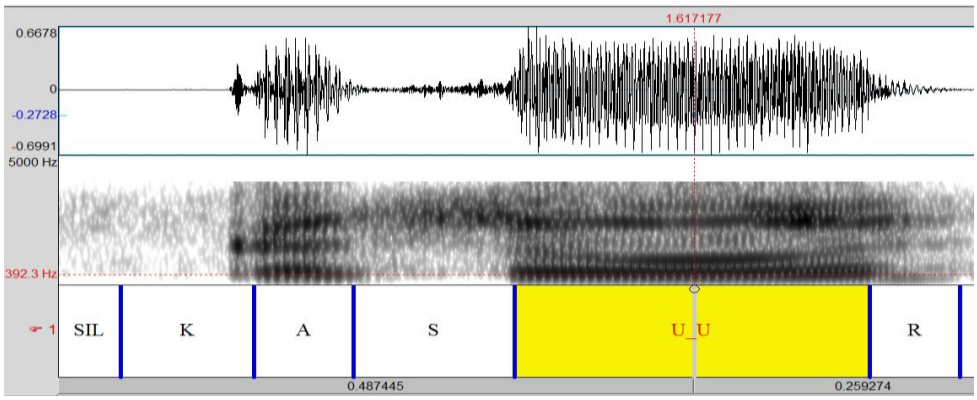
Table 41: *Difference among /a:/ Vowels*

Accents	**F1**	**F2**
Punjabi-Pashto	0.002	---
Urdu- Sindhi	---	0.003
Punjabi-Sindhi	---	0.000
Sindhi-Pashto	---	0.000
Sindhi-Saraiki	---	0.005

The result shows after analyzing F1, F2 values of A_A vowels among the utterances of the speakers of all six major languages. A_A vowels uttered by the speakers of Sindhi language are significantly different from Urdu, Punjabi, Pashto and Saraiki A_A vowels. A_A vowels uttered by the Punjabi speakers are also different from Pashto A_A vowel.

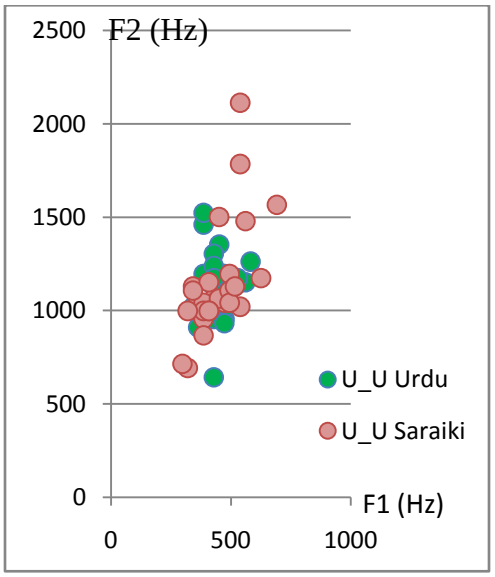
4.4. Acoustic Analysis of U_U (u:) Vowel

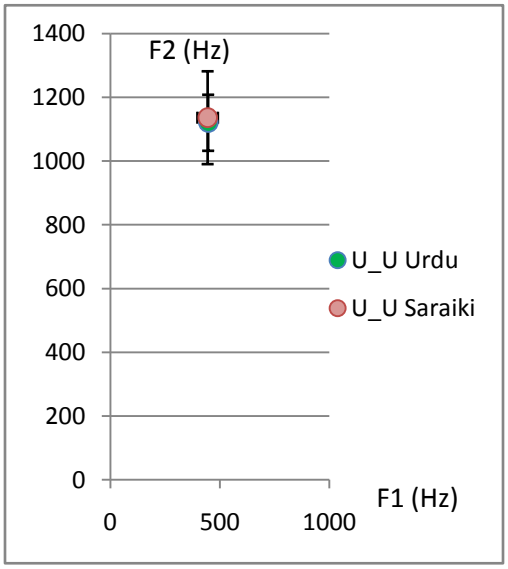
/u:/ vowel is produced when the tongue pushes forward by minimizing the gap between the tongue body and the palate but lips are close and in rounded posture (Skandera & Burleigh, 2005). In Urdu language, /u:/ vowel has slightly up F1 and lower F2 values (300-700 Hz) than /i:/ vowel.

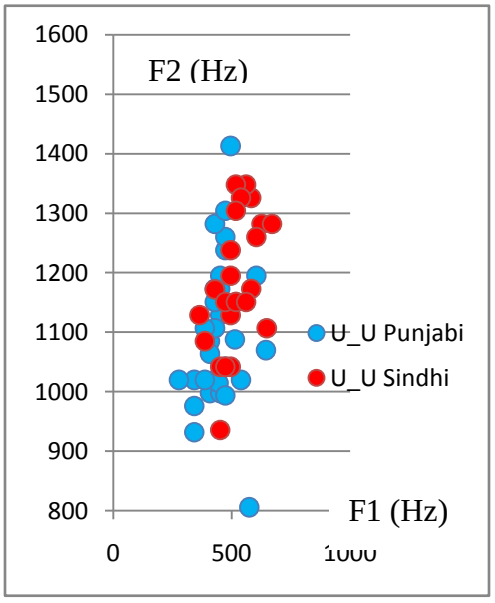


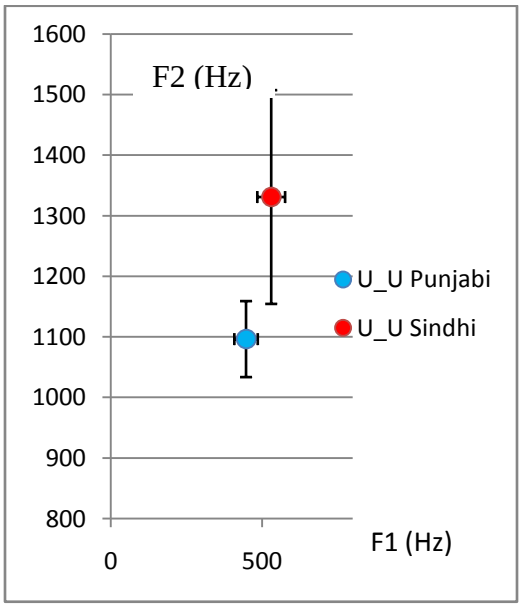
The recorded data was processed by PRAAT software. The data is based on the F1, F2 values of /u:/ vowels. For the understanding of the data variation cluster graphs have been made which show the similar as well as contrastive behavior of the different vowels uttered by the speakers of six languages. These graphical representations show that frequencies of vowels vary from each other but some pair wise comparisons show more contrastive behavior than the other. Few of them are elaborated here;

Pair wise Comparison of U_U Vowels Uttered by Different Language Speakers









The pair wise comparison shows rate of variation of the formant frequencies (F1, F2) among the utterances from the speakers of the different languages. The pair wise comparison of U_U vowels uttered by Urdu and Saraiki speakers show less variation than the U_U vowels uttered by Punjabi and Sindhi speakers.

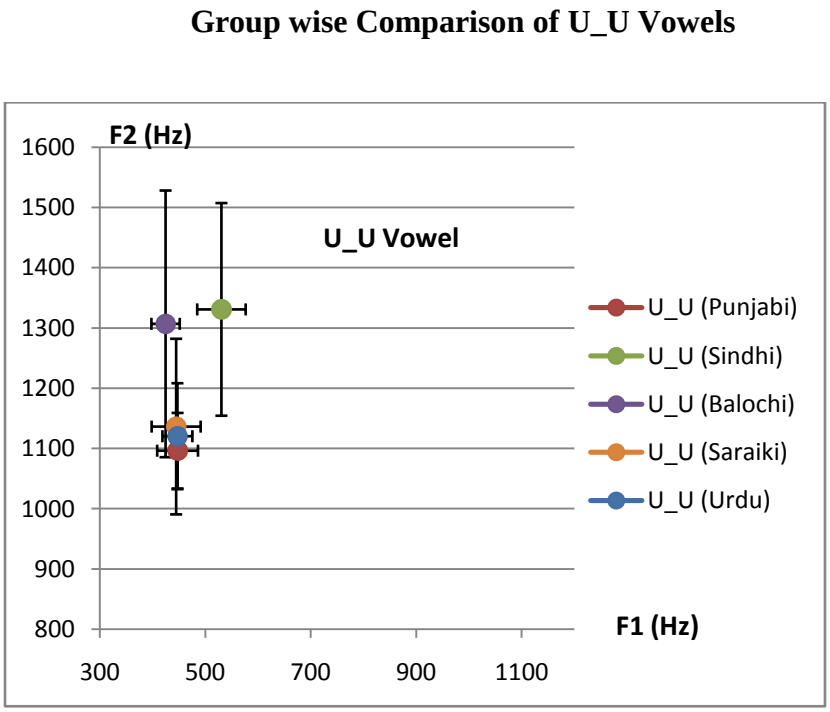


Figure 28: Mean & St. Deviation of /u:/ Vowels

Figure 28: shows the group wise comparison of the mean and standard deviation of (the long vowel /u:/) U_U vowels uttered by the speakers of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages. This group wise comparison shows that the formant frequency values of U_U vowel uttered by the Balochi and Sindhi speakers are drastically different from the formant frequencies of I_I vowel uttered by the speakers of Urdu, Punjabi.

ANOVA test has been run for the verification of the results.

One way ANOVA test was run to determine, if there was any significant difference in U_U vowel (F1, F2) uttered by the speakers of the six languages. There were no outliers and the data was normally distributed for each group and assessed by the Tukey *post-hoc* test [p < .05]. There is statistically significant difference among five groups as determined by one way ANOVA [F (5, 167) = 8.53, *P* = .000]. A Tukey *post-hoc* test revealed that F1 of U_U (u:) vowel uttered by the speakers of Urdu language are significantly different from Sindhi U_U [*P* = .000]. U_U vowel uttered by the speakers of Punjabi language are significantly different from U_U vowels uttered by Sindhi speakers [*P* = .000]. U_U vowels uttered by the speakers of Sindhi language are significantly different from the vowels uttered by Balochi [*P* = .000] and Saraiki [*P* = .000] speakers (For details see APPENDIX E).

Statistical Analysis of U_U (u:) Vowel

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
U_U.F1	Between Groups	228891.193	5	45778.239	8.538	.000
	Within Groups	895406.101	167	5361.713		
	Total	1124297.295	172			
U_U.F2	Between Groups	18871697.895	5	3774339.579	35.352	.000
	Within Groups	17829510.833	167	106763.538		
	Total	36701208.728	172			

Multiple Comparisons of U_U (u:) Vowel

Dependent Variable	(I) Group	(J) Group	Sig.	Dependent Variable	(I) Group	(J) Group	Sig.
U_U.F1	Urdu	Punjabi	1.000	U_U.F2	Urdu	Punjabi	1.000
		Sindhi	.000			Sindhi	.132
		Balochi	.878			Balochi	.314
		Pashto	.815			Pashto	.000
		Saraiki	1.000			Saraiki	1.000
	Punjabi	Urdu	1.000		Punjabi	Urdu	1.000
		Sindhi	.000			Sindhi	.066
		Balochi	.871			Balochi	.190
		Pashto	.806			Pashto	.000
		Saraiki	1.000			Saraiki	.997
	Sindhi	Urdu	.000		Sindhi	Urdu	.132
		Punjabi	.000			Punjabi	.066
		Balochi	.000			Balochi	1.000
		Pashto	.000			Pashto	.000
		Saraiki	.000			Saraiki	.198
	Balochi	Urdu	.878		Balochi	Urdu	.314
		Punjabi	.871			Punjabi	.190
		Sindhi	.000			Sindhi	1.000
		Pashto	1.000			Pashto	.000
		Saraiki	.921			Saraiki	.417
	Pashto	Urdu	.815		Pashto	Urdu	.000
		Punjabi	.806			Punjabi	.000
		Sindhi	.000			Sindhi	.000
		Balochi	1.000			Balochi	.000
		Saraiki	.874			Saraiki	.000
	Saraiki	Urdu	1.000		Saraiki	Urdu	1.000
		Punjabi	1.000			Punjabi	.997
		Sindhi	.000			Sindhi	.198
		Balochi	.921			Balochi	.417
		Pashto	.874			Pashto	.000

Table 42: *Difference among /u:/ Vowels*

Accents	**F1**
Urdu-Sindhi	.000
Punjabi-Sindhi	.000
Sindhi-Balochi	.000
Sindhi-Saraiki	.000

Pashto U_U vowel has not been used in this comparison. Because Pashto U_U vowel is pronounced with more lips rounding and has become the cause of the lowering F3 formants. Therefore, F3 merged with the F2 and it was difficult to find and measure the F2 values of Pashto /u:/ vowels.

The result shows after analyzing F1, F2 values of U_U vowels uttered by the speakers of Sindhi language is significantly different from the formant values of U_U vowels pronounced by the speakers of Urdu, Punjabi, Balochi and Saraiki U_U languages.

After calculating, the mean values of formants (F1, F2) and standard deviation of all these corner vowels. These values have been compared with each other for finding the difference among them. Then the four corner vowels have been mapped with the help of these values mentioned in the table 44.

Table 44: *Mean and Standard Deviation of F1, F2 among Six Languages*

	I_I		**A_E**		**A_A**		**U_U**	
	F1	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
Urdu Mean	417	2302	563	1732	684	1385	447	1120
Urdu Std. Dev.	28	194	35	225	55	105	28	87
Punjabi Mean	412	2364	566	1653	750	1304	447	1096
Punjabi Std. Dev.	24	166	53	173	65	58	38	62
Sindhi Mean	470	2496	521	1936	686	1595	530	1331
Sindhi Std. Dev.	47	209	34	247	76	137	45	176
Balochi Mean	403	2163	501	1800	667	1400	424	1307
Balochi Std. Dev.	39	142	50	141	63	104	26	221
Pashto Mean	427	2183	508	1943	629	1355	423	2032
Pashto Std. Dev.	24	135	36	158	39	54	24	229
Saraiki Mean	446	2271	500	1919	639	1394	444	1136
Saraiki Std. Dev.	36	233	41	174	57	156	46	145

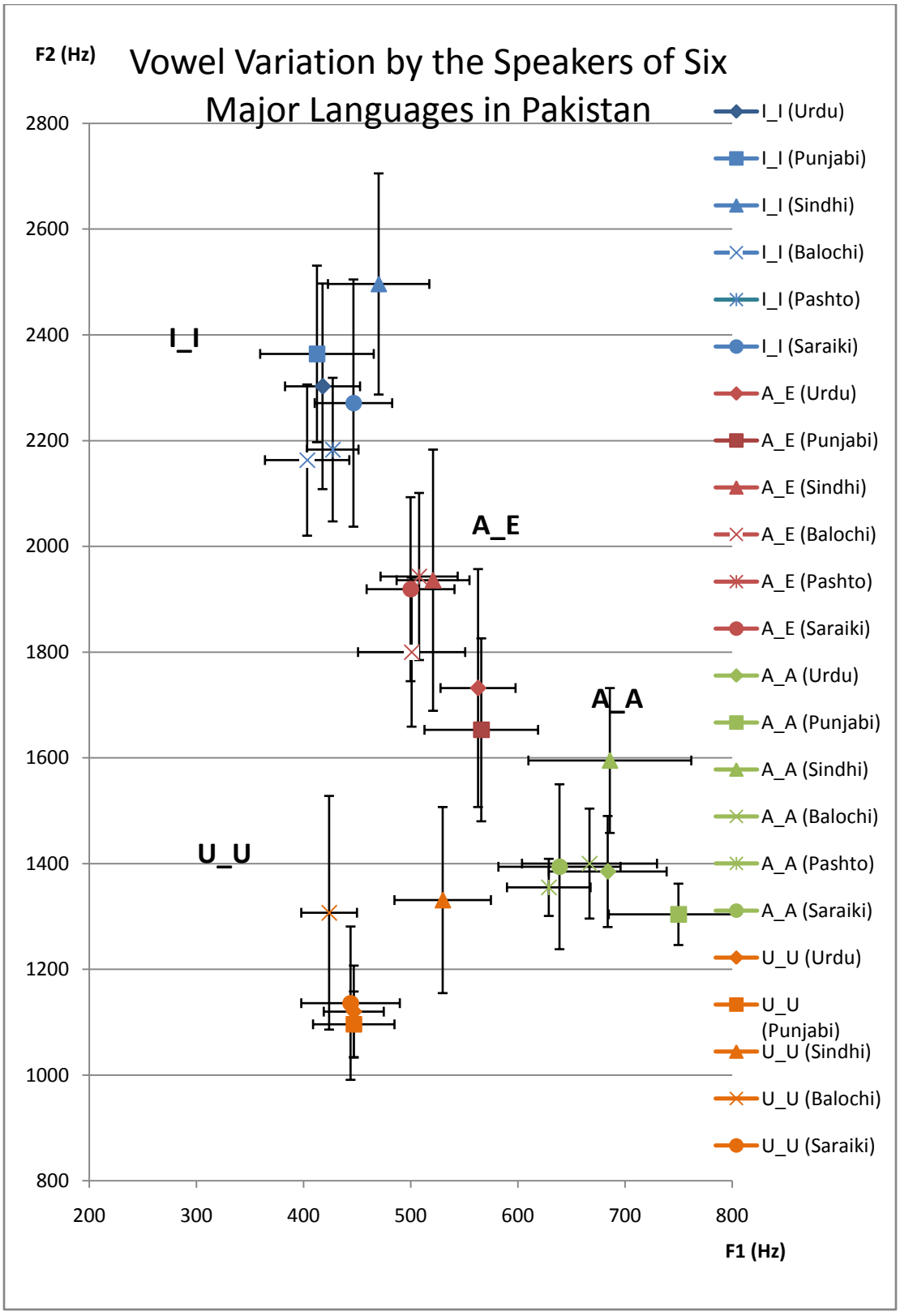


Figure 29: Mean &Std. Deviation of F1, F2 among Corner Vowels by the Speakers of Six Languages

All the six languages; Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki have also been compared pair wise for finding out vowel variations. Accent variations among all these six accents have been calculated by knowing the significant gaps among the values of F1, F2 or both of them because the acoustic properties of a vowel is dependent on first two formants (Reetz & Jongman, 2008).

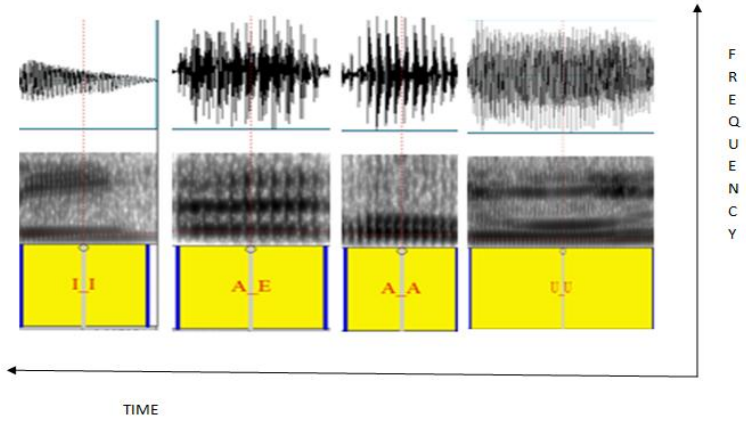


Figure 29: Comparison of Urdu Corner Vowels

Table 43: *Difference among the Vowels of all six Languages*

	Urdu	Punjabi	Sindhi	Balochi	Pashto	Saraiki
Urdu			**i: a: u:**		**u:**	
Punjabi			**i: a: u:**		**æ: a: u:**	
Sindhi				**i:i: u:**	**i: a: u: u:**	**a: u:**
Balochi					**u:**	
Pashto						**u:**
Saraiki						

Red Color for difference based on F1 of vowels

Green Color for difference based on F2 of vowels

Grey shaded area shows repetition of comparisons among languages

It has been cleared by the analysis that vowels are produced with the loud acoustic signals in all six accents of Urdu i.e. Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. First two formant frequencies are the primary acoustic properties of the vowels in all six accents. The dark bands represent the formant frequencies in order to reflect the resonance in vocal tract. Formant frequencies of vowels are directly affected by the ***positions of articulators*** in the oral tract among all accents and the same formant frequencies can be generated with the different articulatory positions. Among the utterances of all six language speakers, the ***locations of formant*** frequencies show the clear difference in each corner vowel. For example, in the utterances of all six languages, /i:/ vowel has first formant frequency around 200-300 Hz and second frequency around 2200-2400 Hz (for the formant values of other vowels see figure 44). In the utterances of the speakers of all the six langauges, ***Vowel height*** is also inversely related with the first formant i.e. the higher the vowel lower will be the F1 value. This is applicable for both front and back vowels. For example, in all six accents, vowels of comparable height /i:/ and /u:/ show comparable first formant frequencies and on the other side /æ:/ and /a:/ show comparable similarity in first formant frequencies. ***Vowel backness*** is reflected in the second formant frequency (Reetz & Jongman, 2008). Larger gap between F2 and F1 reflects the front vowels i.e. /i:/, /æ:/ and smaller gap reflects the back vowels i.e. /a:/ and /u:/ among the utterances by the speakers of all six languages in Pakistan. ***Lips rounding*** lower all the formants of vowels. Among all the utterances by the speakers of six languages, ***duration*** also plays an important role in vowel's identity i.e. long /i:/ differs from short /ɪ/ in terms of duration of the dark bands in spectrum. In all the utterances by the speakers (of Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki languages); the

formant frequencies of the vowels are ***stable in the middle of the vowel*** because at this point vowel is away from the influence of the surrounding consonantal segments (Reetz & Jongman, 2008). This is clear that a vowel has different acoustic properties which are responsible for providing unique information about the particular identity of a specific accent.

After analyzing the data, it is concluded;

- The formant frequencies of the vowels uttered by the speakers of Urdu and Punjabi are not identical but similar to each other.
- The formant frequencies of the vowels uttered by the speakers of Urdu are different from the vowel formant values (utterances) by the speakers of Sindhi, Pashto and Saraiki languages.
- The formant frequencies of the vowels uttered by the speakers of Punjabi language are different from the utterances Sindhi, Balochi, Pashto and Saraiki accents.
- Pair wise comparison of the utterances by the speakers of all six languages shows;
    - The formant frequencies of the vowels uttered by the speakers of Sindhi accent are different from Urdu, Punjabi, Balochi, Pashto and Saraiki speakers' utterances.
    - The formant frequencies of the vowels uttered by the speakers of Balochi accent are different from Urdu, Punjabi, Sindhi, Pashto and Saraiki languages.

- The formant frequencies of the vowels uttered by the speakers of Pashto are different from the formant frequencies (utterances) Urdu, Punjabi, Sindhi, Balochi and Saraiki language speakers.
- The formant frequencies of the vowels uttered by the speakers of Saraiki are also different from the formant values (utterances) of Urdu, Punjabi, Sindhi, Balochi and Pashto language speakers.

CHAPTER 5

CONCLUSION

The present research has been conducted for finding out the accent variation in Pakistani Urdu language. Although Urdu is an official language but Pakistan is a multilingual country where majority of speakers having different first languages. So we have analyzed the affect of the accents of majorly spoken languages on Urdu. Data has been collected by using the medium of telephone or mobile phone. Students of different Pakistani universities have been volunteered for recordings of speech corpus (the list of district names). The quantitative methodology has been adopted for finding out accent variations by means of measuring vowel frequencies.

The present research is based on the acoustic analysis of the spectrograms of the words and their segments. As, this is the first attempt for finding out the accent variation therefore, only deals with the analysis of the corner vowels only. Four corner vowels have been analyzed acoustically by measuring first two formant frequencies F1 and F2 after taking the spectral slices. The formant frequencies of Urdu vowels have been extracted from the wave files uttered by the speakers of six majorly spoken languages; Urdu, Punjabi, Sindhi, Balochi, Pashto and Saraiki. The first two frequencies (F1, F2) have been measured in Hertz of the “corner vowels” /i:/, /æ:/, /a:/ and /u:/ vowels uttered by the speakers of all the six languages. Mean values of formants have been calculated from the utterances of the speakers of each language then compared pair wise and group wise.

The acoustic phonetic analysis of the corner vowels has shown the idiosyncratic behavior of /i:/, /æ:/, /a:/ and /u:/ vowels among utterances by the speakers of six languages. The analysis also verified that the formant frequencies (F1, F2) of the vowels uttered by the speakers of all languages are different from each other. Even the formant frequencies of Urdu and Punjabi speakers' utterances are similar but not identical to each other. So, we can say that the vowels uttered by Urdu and Punjabi speakers can map on each other but are different from Sindhi, Balochi, Pashto and Saraiki speakers' vowel utterances. The vowels uttered by the speakers of Sindhi, Balochi, Pashto and Saraiki languages are not only different from each other but also show difference from Urdu and Punjabi languages. This study also confirms that accent variation depends on the multilingualism of the people in Pakistan. The other reason is the fact that Urdu is the second or third language to almost all Pakistanis (Rehman, 2002) and the reason of variations among these accents is the affect and influence of the native languages of Pakistani people.

5.1. Recommendations and Future Directions

The present study gives the sufficient information about the vowel variation with reference to formant values. The experimentation shows that two dimensional features F1 and F2 are not enough to identify the different accents spoken in different geographical regions of Pakistan. Therefore it is essential to find out more dimensions of speech data;

- The acoustic analysis of Urdu consonantal inventory and the other vowels are still in need to study in future.

- Supra segmental features of phonetic inventory can be analyzed acoustically by measuring the duration, stress, pitch, intensity, intonation and loudness in all six accents of Pakistani Urdu.

Bibliography

Abbas, S. N. (2012). *Nasal and Nasality in Saraiki.* Retrieved from http://connection.ebscohost.com/c/articles/88819914/nasals-nasality-saraiki

Ahangar, A. A., Oostendrop, v. M., & Soohani, B. (2013). Relationship between Phonological and Geographical Distance: Persian, Sarawani Balochi and Sistani Dialect. *GEMA online Journal of Language Studies , 13(2)*, 117-134.

Andrews, E., Frigau, L., Voyvodic, J., Wright, J., & Voyvodic-Casabo, C. (2013). Brain Science (2076-3425). *Acedamic Journal , 3* (2), 849.

Awan, M. S., Baseer, A., & Sheeraz, M. (2012). Outlining Saraiki Phonetics: A comparative Study of Saraiki and English Sound System. *Language in India: Strength for Today and Bright Hope for Tomorrow , 12* (1930-2940), 1930-2940.

Bakst, S. (2012). Rhotics and Retroflex in Indic and Darvidian.

Benzeghiba, M., Mori, D. R., Deroo, O., Dupont, S., Jouvet, D., Fissore, L., et al. (2006). Impact of Variabilities on Speech Recognition. *Multimedia Communications Eurecom.* Saint-Petersburg, RUSSIAN FEDERATION.

Bhatia, T. K., & William, R. C. (Eds.). (2013). *The Handbook of Bilingualism and Multilingualism* (2nd ed., Vol. 13). Blackwell Handbooks in Linguistics.

Cole, J. (2012). Prosody in Context: A Review. (61801), 55.

Din, R. U., & Rahman, G. (2011). The Acoustic Analysis of Pashto Vowels. *Language in India: Strength for Today and Bright Future for Tomorrow , 11*, 793-797.

Elfenbein, J. (1997). Phonologies of Asia and Africa: Including the Caucasus. Johannes Gutenberg Universitat Mainz.

Ghai, W., & Singh, N. (2013). Phone Based Acoustic Modeling for Automatic Speech Recognition for Punjabi Language. *Journal of Speech Sciences , 3*, 69-83.

Gordon, R. G. (Ed.). (2005). *Ethnologue: Languages of the World, 15th Edition* (15th ed.).

Gorter, D., Cenoz, J., Nunes, P., Riganti, P., Onofri, L., Puzzo, B., et al. (2003). "Cultural Diversity as an Asset for Human Welfare and Development", Benefits of Linguistic Diversity and Multilingualism. *Sustainable Development in a Diverse World* .

Haq, A. M. (1967). Multani Zuban Aur Uska Urdu Se Taalluq.

Hillenbrand, J., Getty, A., Clark, M. J., & Michael, W. (1995). Acoustic Characteristics of American English Vowels. *Western Michigan University , 97* (49008), 3099-3111.

Hussain, S. (1997). *Phonetic Correlates of Lexical Stress in Urdu.* United States: Northwestern University.

Hussain, S. (2006). *Phonological Processing for Urdu Text to Speech System.* Retrieved from http://www.researchgate.net/publication/228410903_Phonological_Processing_for_Urdu_Text_to_Speech_System

Ijaz, M. (2002-2003). *Phonemic Inventory of Pashto.* CRULP, Annual Student Report 2002-2003, National University of Computer and Emerging Sciences, Center for Research in Urdu Language Processing (CRULP), B Block, Faisal Town, Lahore.

Karamat, N. (2012, April-May). *Phonemic Inventory of Punjabi.* Retrieved from Center for Language Engineering and Technology: http://www.cle.org.pk/information/people/nayyarakaramat.html

Keerio, A. (2010, November). *Acoustic Analysis of Sindhi Speech: A Pre-curser for an ASR System.*

Ladefoged, P. (1993). *A Course in Phonetics* (3rd ed. ed.). Fort Worth: Harcourt Brace College Publishers .

Ladefoged, P. (2005). *A Course in Phonetics* (5th ed.). Fort Worth: Harcourt Brace College Publishers.

Lata, S. (2011). *Challenges for Design of Pronunciation Lexicon Specification (PLS) for Punjabi Language.* Retrieved from http:www.mit.gov.in/WSI/papers/871621ltc-111-swaranlata.pdf

Latif, A. (2006). *Phonemic Inventory of Saraiki Language and Acoustic Analysis of Voiced Implosives.* University of Engineering and Technology, Center for Language Engineering, Lahore.

Mahmood, M. S. (2004, July 31). *CSS Forum.* Retrieved March 1, 2013, from http://www.mofa.gov.pk/Publications/Constitution.pdf

McCombs, C. J. (2006). *The Acoustic Properties of Vowels: A Tool for Improving Articulation and Comprehension of English.* Forum on Public Policy: A Journal of the Oxford Round Table.

Mughal, M. A. (2011, June 6). *The National Language of Pakistan, Urdu.* Retrieved December 22, 2012, from Retrieved from http://www.ssrn.com/author/1697634

Nasir, S. (2012). Nasal and Nasality in Saraiki. *Language in india* (88819914).

Oxford Urdu English Dictionary (1st ed.). (2013). Pakistan: Oxford Publication.

Qandeel, H., Mahmood, R., & Mahmood, A. M. (2012). Phonological Make-up of English Loanwords Incorporated into Punjabi via Urdu. *Journal of Language Teaching and Research (JLTR) , 3*, 838-843.

Raza, A. A. (2009, July). Design and Development of an Automatic Speech Recognition System for Urdu. 211. Lahore, Pakistan: NUCES-FAST.

Reetz, H., & Jongman, A. (2008). Acoustic Characteristics of Speech Sounds. In H. Reetz, & A. Jongman, *Phonetics: Transcription, Production, Acoustics and Perception* (1st ed.).

Rehman, D. T. (2006). Urdu as an Islamic Language-the Annual of Urdu Studies. (M. U. Memon, Ed.) *Urdu Studies , 22*, pp. 101-119.

Rehman, T. (2002). Retrieved from http://www.sil.org/asia/ldc/parallel_papers/tariq_rahman.pdf

Rehman, T. (2002). *Language Ideology and Power: Language Learning among the Muslims of Pakistan and North India* (illustrated ed., Vol. xix). the University of Michigan: Oxford University Press.

Rehman, T. (2002). *Language policy, multilingualism and language vitality in Pakistan.* Retrieved from http://www.sil.org/asia/ldc/parallel_papers/tariq_rahman.pdf

Roach, P. (2009). *English Phonetics and Phonology: A Practical Course* (4th ed.). Cambridge University Press.

Robert, T. (2000, June). Clitics and Agreement. 213. Massachusetts Institute of Technology.

Rovira, M. i. (2007). Multilingualism, An Emerging Value. *Noves SL. Revista de Sociolinguistica* , 1-5.

Sabir, D. A. (2004). Few Examples of Balochi and Brahui Morphological Similarities. *Proceedings of the International Symposium on "Linguistic contacts in Balochistan ancient and modern time"* , pp. 60-151.

Saleem, A. M., Kabir, H., Riaz, M. K., Rafique, M. M., Khalid, N., & Shahid, S. R. (2012). *Urdu Consonantal and Vocalic Sounds.* Annual Report, Center for Research in Urdu Language Processing (CRULP), Lahore.

Skandera, P., & Burleigh, P. (2005). *A Manual of English Phonetics and Phonology: Twelve Lessons with an Integrated Course in Phonetic Transcription.* Narr Francke Attempto Verlag GmbH + Co. KG Dischingerweg 5. D-72070 Tubingen.

Stantic, D., & Jo, J. (2012). Accent Identification by Clustering and Scoring Formants. *International Journal of Computer and Communication Engineering , 63*, 232-237.

UCL, Psychology & Language Sciences. (2004). (UCL Division of Psychology and Language Sciences, Mark Huckvale University College London) Retrieved April 26, 2015, from Speech, Hearing & Phonetic Sciences: http://www.phon.ucl.ac.uk/home/mark/accent/

Universitat Bielefeld. (2010, November 9). (Germany) Retrieved August 12, 2013, from http://www.uni-bielefeld.de/(en)/lili/personen/dschlangen/publications.html

Yan, Q., & Vaseghi, S. (2002). *Brunel University.* Retrieved from http://www.dea.brunel.ac.uk/cmsp/publicationlist.htm

Zahid, S. (2010). An Acoustic Analysis of Vowel Nazalaization in Punjabi Speakers' Production of English in Pakistan. Faisalabad, Pakistan: Government College University, Faisalabad, Pakistan.

Zia, W. (Ed.). (2011, Febrary 12). *Pakistan 6th Population and Housing Census-Pakistan*. Retrieved from http://www.paknetmag.blogspot.com/2011/..../2011-pakistan-6th-population

APPENDIX A

Consent Form

Dear Respondent!

I am Mahwish Farooq, M. Phil scholar in University of Management and Technology. I am doing research on the topic "An Acoustic Phonetic Study of Six Accents of Urdu in Pakistan" I want to have your consent in my research, I assure you for confidentiality of the information given by you. This information will be used for research purpose only. I would be grateful for your valued contribution.

Thanks in Anticipation

Mahwish Farooq

APPENDIX B

Sr. no.	Subdivisions	Major Languages	Districts
1	Balochistan	Balochi & 10-12 minor languages are used	30
2	Khyber Pakhtunkhwa	Pashto & Almost 35 minor languages are used	24
3	Punjab	Punjabi & Almost 30 minor languages are used	36
4	Sindh	Sindhi & Almost 31-34 minor languages are used	23
5	Islamabad Capital Territory	Udru, Punjabi, English	1
6	Federally Administered Tribal Areas (FATA)	Pashto, Urdu, Hindko	7 tribal agencies and 6 frontier regions
7	Azad Kashmir	Kashmiri, Urdu, Hindi	10
8	Gilgit–Baltistan	Balti, Burushaski, Chainies, Domaaki, Gojal, Shina, Wakhi, Urdu	7

APPENDIX C
Guidelines for Segment Marking

For the purpose of segment labeling we are using the Case Insensitive Speech Assessment Method Phonetic Alphabet (CISAMPA). Since the IPA symbols are difficult to use for our applications, symbols in the Speech Assessment Method Phonetic Alphabet (SAMPA) are matched to the IPA symbols and used for labeling. Some additions to the existing SAMPA symbols have been made.

The following table shows Urdu letters, their corresponding IPA symbols and the CISAMPA.

Urdu Letter	**IPA**	**CISAMPA**	**Urdu Letter**	**IPA**	**CISAMPA**	**Urdu Letter**	**IPA**	**CISAMPA**
پ	p	P	خ	x	X	ق	q	Q
پھ	p^h	P_H	غ	ɣ	G_G	ع	ʔ	Y
ب	b	B	ہ،ح	h	H	ف	f	F
بھ	b^h	B_H	ل	l	L	و	v	V
م	m	M	لھ	l^h	L_H	س،ص،ث	s	S
مھ	m^h	M_H	ر	r	R	ذ،ز،ظ،ض	z	Z
ت،ط	ʈ	T_D	رھ	r^h	R_H	ش	ʃ	S_H

تھ	ʈʰ	T_D_H	ڑ	ɽ	R_R	ژ	ʒ	Z_Z
د	ɖ	D_D	ڑھ	ɽʰ	R_R_H	ُ	ʊ	U
دھ	ɖʰ	D_D_H	ی	j	J	ء،َ	ə	A
ٹ	t	T	یھ	jʰ	J_H	ں ِ	ĩ	I_N
ٹھ	tʰ	T_H	چ	ʧ	T_S	آ،ا	ɑ:	A_A
ڈ	d	D	چھ	ʧʰ	T_S_H	آں،اں	ɑ̃:	A_A_N
ڈھ	dʰ	D_H	ج	ʤ	D_Z	ی	i:	I_I
ن	n	N	جھ	ʤʰ	D_Z_H	ِیں	ĩ:	I_I_N
نھ	nʰ	N_H	وُ	u:	U_U	ے	e:	A_Y
ک	k	K	وُں	ũ:	U_U_N	یں	ẽ:	A_Y_N
کھ	kʰ	K_H	و	o:	O_O	ہ ِ	E	A_Y_H
گ	g	G	وں	õ:	O_O_N		E	A_E_H
گھ	gʰ	G_H	وَ	ɔ:	O	ں ُ	õ	U_N
ن in نگ , نک ,نکھ ,نگھ	ŋ	N_G	وَں	ɔ̃:	O_N	ں َ	ə̃	A_N

APPENDIX D
Comparison of Oral Vowels in Six Major Languages

	Urdu Vowels	**Punjabi Vowels**	**Sindhi Vowels**	**Balochi Vowels**	**Pashto Vowels**	**Saraiki Vowels**
1	iː	iː	I	iː	iː	iː
	ɪ	Ɪ	ɪ		i	i
2	eː	e		e	e	eː
						e
3	æː			æ	æ	æ
4		ɛː	ɛː			
			ɛ			ɛ
5	ɑː	aː	ɑ	ɑː	a	aː
				ɑ		
6	ɔː	ɔː				
			ɔ			ɔ
7	oː		oː	oː		oː
		o		o	o	
8	uː	uː	u	uː	u	uː
	ʊ	ʊ	ʊ		ʊ	
						u
9	ə	ə	ə	ə	Ə	
						a
10						ʌ

APPENDIX E

Oneway

Descriptive

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
	Urdu	30	684.5667	111.30470	20.32137	643.0048	726.1285	386.00	867.00
	Punjabi	30	750.9000	131.64565	24.03510	701.7427	800.0573	451.00	998.00
	Sindhi	30	686.6333	153.32375	27.99296	629.3813	743.8854	386.00	977.00
A_AF1	Balochi	30	667.4333	126.47384	23.09086	620.2072	714.6594	429.00	910.00
	Pushto	30	629.8333	79.67784	14.54712	600.0811	659.5855	495.00	801.00
	Saraiki	30	639.1000	114.62032	20.92671	596.3001	681.8999	386.00	823.00
	Total	180	676.4111	126.25746	9.41068	657.8410	694.9812	386.00	998.00
	Urdu	30	1385.0667	211.16490	38.55326	1306.2164	1463.9169	932.00	1894.00
	Punjabi	30	1304.0667	117.96374	21.53713	1260.0183	1348.1151	1085.00	1566.00
	Sindhi	30	1596.7333	275.45786	50.29149	1493.8757	1699.5910	1020.00	2069.00
A_AF2	Balochi	30	1400.6667	208.44438	38.05656	1322.8323	1478.5011	932.00	1664.00
	Pushto	30	1355.8000	109.81282	20.04899	1314.7952	1396.8048	1195.00	1676.00
	Saraiki	30	1392.1667	313.15162	57.17340	1275.2339	1509.0994	729.00	2069.00
	Total	180	1405.7500	234.56976	17.48380	1371.2491	1440.2509	729.00	2069.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	280854.311	5	56170.862	3.799	.003
A_AF1	Within Groups	2572575.267	174	14784.915		
	Total	2853429.578	179			
	Between Groups	1498418.517	5	299683.703	6.244	.000
A_AF2	Within Groups	8350693.233	174	47992.490		
	Total	9849111.750	179			

Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
A_AF1	Urdu	Punjabi	-66.33333	31.39524	.286	-156.8063	24.1397
		Sindhi	-2.06667	31.39524	1.000	-92.5397	88.4063
		Balochi	17.13333	31.39524	.994	-73.3397	107.6063
		Pushto	54.73333	31.39524	.505	-35.7397	145.2063
		Saraiki	45.46667	31.39524	.697	-45.0063	135.9397
	Punjabi	Urdu	66.33333	31.39524	.286	-24.1397	156.8063
		Sindhi	64.26667	31.39524	.320	-26.2063	154.7397
		Balochi	83.46667	31.39524	.089	-7.0063	173.9397
		Pushto	121.06667*	31.39524	.002	30.5937	211.5397
		Saraiki	111.80000*	31.39524	.006	21.3270	202.2730
	Sindhi	Urdu	2.06667	31.39524	1.000	-88.4063	92.5397
		Punjabi	-64.26667	31.39524	.320	-154.7397	26.2063
		Balochi	19.20000	31.39524	.990	-71.2730	109.6730
		Pushto	56.80000	31.39524	.463	-33.6730	147.2730
		Saraiki	47.53333	31.39524	.656	-42.9397	138.0063
	Balochi	Urdu	-17.13333	31.39524	.994	-107.6063	73.3397
		Punjabi	-83.46667	31.39524	.089	-173.9397	7.0063
		Sindhi	-19.20000	31.39524	.990	-109.6730	71.2730
		Pushto	37.60000	31.39524	.838	-52.8730	128.0730
		Saraiki	28.33333	31.39524	.945	-62.1397	118.8063
	Pushto	Urdu	-54.73333	31.39524	.505	-145.2063	35.7397
		Punjabi	-121.06667*	31.39524	.002	-211.5397	-30.5937
		Sindhi	-56.80000	31.39524	.463	-147.2730	33.6730
		Balochi	-37.60000	31.39524	.838	-128.0730	52.8730
		Saraiki	-9.26667	31.39524	1.000	-99.7397	81.2063
	Saraiki	Urdu	-45.46667	31.39524	.697	-135.9397	45.0063
		Punjabi	-111.80000*	31.39524	.006	-202.2730	-21.3270
		Sindhi	-47.53333	31.39524	.656	-138.0063	42.9397
		Balochi	-28.33333	31.39524	.945	-118.8063	62.1397

		Pushto	9.26667	31.39524	1.000	-81.2063	99.7397
		Punjabi	81.00000	56.56412	.707	-82.0033	244.0033
		Sindhi	-211.66667*	56.56412	.003	-374.6699	-48.6634
	Urdu	Balochi	-15.60000	56.56412	1.000	-178.6033	147.4033
		Pushto	29.26667	56.56412	.995	-133.7366	192.2699
		Saraiki	-7.10000	56.56412	1.000	-170.1033	155.9033
		Urdu	-81.00000	56.56412	.707	-244.0033	82.0033
		Sindhi	-292.66667*	56.56412	.000	-455.6699	-129.6634
	Punjabi	Balochi	-96.60000	56.56412	.529	-259.6033	66.4033
		Pushto	-51.73333	56.56412	.942	-214.7366	111.2699
		Saraiki	-88.10000	56.56412	.628	-251.1033	74.9033
		Urdu	211.66667*	56.56412	.003	48.6634	374.6699
		Punjabi	292.66667*	56.56412	.000	129.6634	455.6699
	Sindhi	Balochi	196.06667*	56.56412	.009	33.0634	359.0699
		Pushto	240.93333*	56.56412	.000	77.9301	403.9366
		Saraiki	204.56667*	56.56412	.005	41.5634	367.5699
A_AF2		Urdu	15.60000	56.56412	1.000	-147.4033	178.6033
		Punjabi	96.60000	56.56412	.529	-66.4033	259.6033
	Balochi	Sindhi	-196.06667*	56.56412	.009	-359.0699	-33.0634
		Pushto	44.86667	56.56412	.968	-118.1366	207.8699
		Saraiki	8.50000	56.56412	1.000	-154.5033	171.5033
		Urdu	-29.26667	56.56412	.995	-192.2699	133.7366
		Punjabi	51.73333	56.56412	.942	-111.2699	214.7366
	Pushto	Sindhi	-240.93333*	56.56412	.000	-403.9366	-77.9301
		Balochi	-44.86667	56.56412	.968	-207.8699	118.1366
		Saraiki	-36.36667	56.56412	.988	-199.3699	126.6366
		Urdu	7.10000	56.56412	1.000	-155.9033	170.1033
		Punjabi	88.10000	56.56412	.628	-74.9033	251.1033
	Saraiki	Sindhi	-204.56667*	56.56412	.005	-367.5699	-41.5634
		Balochi	-8.50000	56.56412	1.000	-171.5033	154.5033
		Pushto	36.36667	56.56412	.988	-126.6366	199.3699

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

A_AF1

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Pushto	30	629.8333	
Saraiki	30	639.1000	
Balochi	30	667.4333	667.4333
Urdu	30	684.5667	684.5667
Sindhi	30	686.6333	686.6333
Punjabi	30		750.9000
Sig.		.463	.089

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 30.000.

A_AF2

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Punjabi	30	1304.0667	
Pushto	30	1355.8000	
Urdu	30	1385.0667	
Saraiki	30	1392.1667	
Balochi	30	1400.6667	
Sindhi	30		1596.7333
Sig.		.529	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 30.000.

Oneway

Descriptive

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
	Urdu	20	563.1000	71.62027	16.01478	529.5807	596.6193	429.00	692.00
	Punjabi	20	566.6000	106.76112	23.87251	516.6343	616.5657	364.00	714.00
	Sindhi	18	521.7778	68.59134	16.16713	487.6681	555.8874	408.00	692.00
A_E.F1	Balochi	10	501.5000	100.58523	31.80784	429.5457	573.4543	364.00	714.00
	Pushto	27	508.3778	72.12132	13.87975	479.8475	536.9080	408.00	735.00
	Saraiki	27	500.7037	83.91746	16.14992	467.5071	533.9003	364.00	670.00
	Total	122	526.6082	86.04382	7.79004	511.1858	542.0306	364.00	735.00
	Urdu	20	1732.1500	450.49216	100.73311	1521.3132	1942.9868	783.00	2528.00
	Punjabi	20	1653.2000	347.36218	77.67254	1490.6295	1815.7705	704.00	2091.00
	Sindhi	18	1936.3333	247.85456	58.41988	1813.0782	2059.5885	1370.00	2464.00
A_E.F2	Balochi	10	1800.5000	283.16005	89.54307	1597.9395	2003.0605	1282.00	2331.00
	Pushto	27	1943.1852	317.96142	61.19170	1817.4038	2068.9665	1413.00	2845.00
	Saraiki	27	1919.4074	349.92969	67.34400	1780.9798	2057.8350	1454.00	2987.00
	Total	122	1843.0820	355.55173	32.19014	1779.3531	1906.8108	704.00	2987.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	92435.684	5	18487.137	2.669	.025
A_E.F1	Within Groups	803392.427	116	6925.797		
	Total	895828.112	121			
	Between Groups	1569726.338	5	313945.268	2.653	.026
A_E.F2	Within Groups	13726734.843	116	118333.921		
	Total	15296461.180	121			

Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
A_E.F1	Urdu	Punjabi	-3.50000	26.31691	1.000	-79.7644	72.7644
		Sindhi	41.32222	27.03805	.647	-37.0320	119.6764
		Balochi	61.60000	32.23150	.401	-31.8044	155.0044
		Pushto	54.72222	24.55200	.233	-16.4276	125.8720
		Saraiki	62.39630	24.55200	.121	-8.7535	133.5461
	Punjabi	Urdu	3.50000	26.31691	1.000	-72.7644	79.7644
		Sindhi	44.82222	27.03805	.562	-33.5320	123.1764
		Balochi	65.10000	32.23150	.338	-28.3044	158.5044
		Pushto	58.22222	24.55200	.175	-12.9276	129.3720
		Saraiki	65.89630	24.55200	.086	-5.2535	137.0461
	Sindhi	Urdu	-41.32222	27.03805	.647	-119.6764	37.0320
		Punjabi	-44.82222	27.03805	.562	-123.1764	33.5320
		Balochi	20.27778	32.82295	.990	-74.8406	115.3961
		Pushto	13.40000	25.32346	.995	-59.9854	86.7854
		Saraiki	21.07407	25.32346	.961	-52.3113	94.4595
	Balochi	Urdu	-61.60000	32.23150	.401	-155.0044	31.8044
		Punjabi	-65.10000	32.23150	.338	-158.5044	28.3044
		Sindhi	-20.27778	32.82295	.990	-115.3961	74.8406
		Pushto	-6.87778	30.80732	1.000	-96.1550	82.3994
		Saraiki	.79630	30.80732	1.000	-88.4809	90.0735
	Pushto	Urdu	-54.72222	24.55200	.233	-125.8720	16.4276
		Punjabi	-58.22222	24.55200	.175	-129.3720	12.9276
		Sindhi	-13.40000	25.32346	.995	-86.7854	59.9854
		Balochi	6.87778	30.80732	1.000	-82.3994	96.1550
		Saraiki	7.67407	22.64999	.999	-57.9638	73.3120
	Saraiki	Urdu	-62.39630	24.55200	.121	-133.5461	8.7535
		Punjabi	-65.89630	24.55200	.086	-137.0461	5.2535

		Sindhi	-21.07407	25.32346	.961	-94.4595	52.3113
		Balochi	-.79630	30.80732	1.000	-90.0735	88.4809
		Pushto	-7.67407	22.64999	.999	-73.3120	57.9638
		Punjabi	78.95000	108.78140	.978	-236.2900	394.1900
	Urdu	Sindhi	-204.18333	111.76226	.453	-528.0617	119.6950
		Balochi	-68.35000	133.22946	.996	-454.4386	317.7386
		Pushto	-211.03519	101.48613	.305	-505.1341	83.0637
		Saraiki	-187.25741	101.48613	.441	-481.3563	106.8415
		Urdu	-78.95000	108.78140	.978	-394.1900	236.2900
		Sindhi	-283.13333	111.76226	.123	-607.0117	40.7450
	Punjabi	Balochi	-147.30000	133.22946	.878	-533.3886	238.7886
		Pushto	-289.98519	101.48613	.056	-584.0841	4.1137
		Saraiki	-266.20741	101.48613	.100	-560.3063	27.8915
		Urdu	204.18333	111.76226	.453	-119.6950	528.0617
		Punjabi	283.13333	111.76226	.123	-40.7450	607.0117
	Sindhi	Balochi	135.83333	135.67424	.917	-257.3401	529.0067
		Pushto	-6.85185	104.67495	1.000	-310.1917	296.4880
		Saraiki	16.92593	104.67495	1.000	-286.4139	320.2658
A_E.F2		Urdu	68.35000	133.22946	.996	-317.7386	454.4386
		Punjabi	147.30000	133.22946	.878	-238.7886	533.3886
	Balochi	Sindhi	-135.83333	135.67424	.917	-529.0067	257.3401
		Pushto	-142.68519	127.34257	.872	-511.7140	226.3437
		Saraiki	-118.90741	127.34257	.937	-487.9363	250.1215
		Urdu	211.03519	101.48613	.305	-83.0637	505.1341
		Punjabi	289.98519	101.48613	.056	-4.1137	584.0841
	Pushto	Sindhi	6.85185	104.67495	1.000	-296.4880	310.1917
		Balochi	142.68519	127.34257	.872	-226.3437	511.7140
		Saraiki	23.77778	93.62412	1.000	-247.5376	295.0932
		Urdu	187.25741	101.48613	.441	-106.8415	481.3563
		Punjabi	266.20741	101.48613	.100	-27.8915	560.3063
	Saraiki	Sindhi	-16.92593	104.67495	1.000	-320.2658	286.4139
		Balochi	118.90741	127.34257	.937	-250.1215	487.9363
		Pushto	-23.77778	93.62412	1.000	-295.0932	247.5376

Homogeneous Subsets

A_E.F1

Tukey HSD

Group	N	Subset for alpha = 0.05
		1
Saraiki	27	500.7037
Balochi	10	501.5000
Pushto	27	508.3778
Sindhi	18	521.7778
Urdu	20	563.1000
Punjabi	20	566.6000
Sig.		.169

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 18.202.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

A_E.F2

Tukey HSD

Group	N	Subset for alpha = 0.05
		1
Punjabi	20	1653.2000
Urdu	20	1732.1500
Balochi	10	1800.5000
Saraiki	27	1919.4074
Sindhi	18	1936.3333
Pushto	27	1943.1852

Sig.		.120

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 18.202.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Oneway

Descriptive

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
	Urdu	30	417.8667	57.78720	10.55045	396.2886	439.4448	277.00	561.00
	Punjabi	30	412.6667	49.85934	9.10303	394.0489	431.2845	320.00	517.00
	Sindhi	30	470.2333	94.61453	17.27417	434.9037	505.5630	320.00	757.00
I_I.F1	Balochi	20	403.5000	78.73373	17.60540	366.6515	440.3485	304.00	648.00
	Pashto	30	427.3667	48.11802	8.78511	409.3991	445.3342	324.00	539.00
	Saraiki	30	446.8000	72.30081	13.20026	419.8024	473.7976	298.00	604.00
	Total	170	431.2824	70.89840	5.43766	420.5479	442.0168	277.00	757.00
	Urdu	30	2302.7000	388.86450	70.99662	2157.4956	2447.9044	1697.00	3205.00
	Punjabi	30	2364.6000	333.93398	60.96772	2239.9070	2489.2930	1807.00	3052.00
	Sindhi	30	2496.3000	418.22772	76.35759	2340.1312	2652.4688	1036.00	3161.00
I_I.F2	Balochi	20	2163.2000	285.97711	63.94642	2029.3586	2297.0414	1676.00	2790.00
	Pashto	30	2183.1667	271.59796	49.58678	2081.7503	2284.5830	1828.00	2812.00
	Saraiki	30	2271.2000	467.77599	85.40382	2096.5296	2445.8704	1346.00	3118.00
	Total	170	2304.7235	383.20206	29.39026	2246.7042	2362.7428	1036.00	3205.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
I_I.F1	Between Groups	84432.180	5	16886.436	3.620	.004
	Within Groups	765060.267	164	4665.002		
	Total	849492.447	169			
I_I.F2	Between Groups	2086300.039	5	417260.008	3.011	.013
	Within Groups	22730305.967	164	138599.427		
	Total	24816606.006	169			

Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
I_I.F1	Urdu	Punjabi	5.20000	17.63520	1.000	-45.6548	56.0548
		Sindhi	-52.36667*	17.63520	.039	-103.2215	-1.5119
		Balochi	14.36667	19.71675	.978	-42.4907	71.2241
		Pashto	-9.50000	17.63520	.994	-60.3548	41.3548
		Saraiki	-28.93333	17.63520	.573	-79.7881	21.9215
	Punjabi	Urdu	-5.20000	17.63520	1.000	-56.0548	45.6548
		Sindhi	-57.56667*	17.63520	.017	-108.4215	-6.7119
		Balochi	9.16667	19.71675	.997	-47.6907	66.0241
		Pashto	-14.70000	17.63520	.961	-65.5548	36.1548
		Saraiki	-34.13333	17.63520	.384	-84.9881	16.7215
	Sindhi	Urdu	52.36667*	17.63520	.039	1.5119	103.2215
		Punjabi	57.56667*	17.63520	.017	6.7119	108.4215
		Balochi	66.73333*	19.71675	.011	9.8759	123.5907
		Pashto	42.86667	17.63520	.152	-7.9881	93.7215
		Saraiki	23.43333	17.63520	.769	-27.4215	74.2881
	Balochi	Urdu	-14.36667	19.71675	.978	-71.2241	42.4907
		Punjabi	-9.16667	19.71675	.997	-66.0241	47.6907
		Sindhi	-66.73333*	19.71675	.011	-123.5907	-9.8759

		Pashto	-23.86667	19.71675	.831	-80.7241	32.9907
		Saraiki	-43.30000	19.71675	.245	-100.1574	13.5574
		Urdu	9.50000	17.63520	.994	-41.3548	60.3548
		Punjabi	14.70000	17.63520	.961	-36.1548	65.5548
	Pashto	Sindhi	-42.86667	17.63520	.152	-93.7215	7.9881
		Balochi	23.86667	19.71675	.831	-32.9907	80.7241
		Saraiki	-19.43333	17.63520	.880	-70.2881	31.4215
		Urdu	28.93333	17.63520	.573	-21.9215	79.7881
		Punjabi	34.13333	17.63520	.384	-16.7215	84.9881
	Saraiki	Sindhi	-23.43333	17.63520	.769	-74.2881	27.4215
		Balochi	43.30000	19.71675	.245	-13.5574	100.1574
		Pashto	19.43333	17.63520	.880	-31.4215	70.2881
		Punjabi	-61.90000	96.12472	.987	-339.0959	215.2959
		Sindhi	-193.60000	96.12472	.339	-470.7959	83.5959
	Urdu	Balochi	139.50000	107.47070	.786	-170.4144	449.4144
		Pashto	119.53333	96.12472	.815	-157.6625	396.7292
		Saraiki	31.50000	96.12472	.999	-245.6959	308.6959
		Urdu	61.90000	96.12472	.987	-215.2959	339.0959
		Sindhi	-131.70000	96.12472	.745	-408.8959	145.4959
	Punjabi	Balochi	201.40000	107.47070	.422	-108.5144	511.3144
		Pashto	181.43333	96.12472	.414	-95.7625	458.6292
		Saraiki	93.40000	96.12472	.926	-183.7959	370.5959
		Urdu	193.60000	96.12472	.339	-83.5959	470.7959
I_I.F2		Punjabi	131.70000	96.12472	.745	-145.4959	408.8959
	Sindhi	Balochi	333.10000*	107.47070	.027	23.1856	643.0144
		Pashto	313.13333*	96.12472	.017	35.9375	590.3292
		Saraiki	225.10000	96.12472	.184	-52.0959	502.2959
		Urdu	-139.50000	107.47070	.786	-449.4144	170.4144
		Punjabi	-201.40000	107.47070	.422	-511.3144	108.5144
	Balochi	Sindhi	-333.10000*	107.47070	.027	-643.0144	-23.1856
		Pashto	-19.96667	107.47070	1.000	-329.8811	289.9477
		Saraiki	-108.00000	107.47070	.916	-417.9144	201.9144
		Urdu	-119.53333	96.12472	.815	-396.7292	157.6625
	Pashto	Punjabi	-181.43333	96.12472	.414	-458.6292	95.7625

		Sindhi	-313.13333*	96.12472	.017	-590.3292	-35.9375
		Balochi	19.96667	107.47070	1.000	-289.9477	329.8811
		Saraiki	-88.03333	96.12472	.942	-365.2292	189.1625
		Urdu	-31.50000	96.12472	.999	-308.6959	245.6959
		Punjabi	-93.40000	96.12472	.926	-370.5959	183.7959
	Saraiki	Sindhi	-225.10000	96.12472	.184	-502.2959	52.0959
		Balochi	108.00000	107.47070	.916	-201.9144	417.9144
		Pashto	88.03333	96.12472	.942	-189.1625	365.2292

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

I_I.F1

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Balochi	20	403.5000	
Punjabi	30	412.6667	
Urdu	30	417.8667	417.8667
Pashto	30	427.3667	427.3667
Saraiki	30	446.8000	446.8000
Sindhi	30		470.2333
Sig.		.177	.054

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 27.692.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

I_I.F2

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Balochi	20	2163.2000	
Pashto	30	2183.1667	

Saraiki	30	2271.2000	2271.2000
Urdu	30	2302.7000	2302.7000
Punjabi	30	2364.6000	2364.6000
Sindhi	30		2496.3000
Sig.		.339	.221

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 27.692.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

language.

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
	Urdu	30	447.2000	56.82302	10.37442	425.9819	468.4181	342.00	582.00
	Punjabi	30	447.5333	77.40522	14.13219	418.6297	476.4369	277.00	644.00
	Sindhi	30	530.7000	91.90781	16.77999	496.3811	565.0189	364.00	757.00
U_U.F1	Balochi	23	424.7391	53.77412	11.21268	401.4855	447.9928	342.00	582.00
	Pashto	30	423.7000	48.67709	8.88718	405.5237	441.8763	324.00	539.00
	Saraiki	30	444.8000	92.98105	16.97594	410.0803	479.5197	298.00	692.00
	Total	173	454.2601	80.84932	6.14686	442.1271	466.3931	277.00	757.00
	Urdu	30	1120.6000	175.83684	32.10327	1054.9414	1186.2586	643.00	1523.00
	Punjabi	30	1096.5000	125.46019	22.90579	1049.6524	1143.3476	806.00	1413.00
	Sindhi	30	1331.0000	352.85808	64.42278	1199.2406	1462.7594	936.00	2571.00
U_U.F2	Balochi	23	1307.0000	442.47650	92.26273	1115.6588	1498.3412	954.00	2335.00
	Pashto	30	2032.3333	458.50178	83.71059	1861.1260	2203.5407	1020.00	2812.00
	Saraiki	30	1136.5333	291.36276	53.19532	1027.7367	1245.3300	692.00	2113.00
	Total	173	1338.5549	461.92978	35.11987	1269.2335	1407.8763	643.00	2812.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
U_U.F1	Between Groups	228891.193	5	45778.239	8.538	.000
	Within Groups	895406.101	167	5361.713		
	Total	1124297.295	172			
U_U.F2	Between Groups	18871697.895	5	3774339.579	35.352	.000
	Within Groups	17829510.833	167	106763.538		
	Total	36701208.728	172			

Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
U_U.F1	Urdu	Punjabi	-.33333	18.90628	1.000	-54.8420	54.1753
		Sindhi	-83.50000*	18.90628	.000	-138.0086	-28.9914
		Balochi	22.46087	20.29388	.878	-36.0483	80.9701
		Pashto	23.50000	18.90628	.815	-31.0086	78.0086
		Saraiki	2.40000	18.90628	1.000	-52.1086	56.9086
	Punjabi	Urdu	.33333	18.90628	1.000	-54.1753	54.8420
		Sindhi	-83.16667*	18.90628	.000	-137.6753	-28.6580
		Balochi	22.79420	20.29388	.871	-35.7150	81.3034
		Pashto	23.83333	18.90628	.806	-30.6753	78.3420
		Saraiki	2.73333	18.90628	1.000	-51.7753	57.2420
	Sindhi	Urdu	83.50000*	18.90628	.000	28.9914	138.0086
		Punjabi	83.16667*	18.90628	.000	28.6580	137.6753
		Balochi	105.96087*	20.29388	.000	47.4517	164.4701
		Pashto	107.00000*	18.90628	.000	52.4914	161.5086
		Saraiki	85.90000*	18.90628	.000	31.3914	140.4086
	Balochi	Urdu	-22.46087	20.29388	.878	-80.9701	36.0483
		Punjabi	-22.79420	20.29388	.871	-81.3034	35.7150

		Sindhi	-105.96087*	20.29388	.000	-164.4701	-47.4517
		Pashto	1.03913	20.29388	1.000	-57.4701	59.5483
		Saraiki	-20.06087	20.29388	.921	-78.5701	38.4483
		Urdu	-23.50000	18.90628	.815	-78.0086	31.0086
		Punjabi	-23.83333	18.90628	.806	-78.3420	30.6753
	Pashto	Sindhi	-107.00000*	18.90628	.000	-161.5086	-52.4914
		Balochi	-1.03913	20.29388	1.000	-59.5483	57.4701
		Saraiki	-21.10000	18.90628	.874	-75.6086	33.4086
		Urdu	-2.40000	18.90628	1.000	-56.9086	52.1086
		Punjabi	-2.73333	18.90628	1.000	-57.2420	51.7753
	Saraiki	Sindhi	-85.90000*	18.90628	.000	-140.4086	-31.3914
		Balochi	20.06087	20.29388	.921	-38.4483	78.5701
		Pashto	21.10000	18.90628	.874	-33.4086	75.6086
		Punjabi	24.10000	84.36569	1.000	-219.1343	267.3343
		Sindhi	-210.40000	84.36569	.132	-453.6343	32.8343
		Balochi	-186.40000	90.55759	.314	-447.4862	74.6862
	Urdu	Pashto	-911.73333*	84.36569	.000	-1154.9677	-668.4990
		Saraiki	-15.93333	84.36569	1.000	-259.1677	227.3010
		Urdu	-24.10000	84.36569	1.000	-267.3343	219.1343
		Sindhi	-234.50000	84.36569	.066	-477.7343	8.7343
		Balochi	-210.50000	90.55759	.190	-471.5862	50.5862
U_U.F2	Punjabi	Pashto	-935.83333*	84.36569	.000	-1179.0677	-692.5990
		Saraiki	-40.03333	84.36569	.997	-283.2677	203.2010
		Urdu	210.40000	84.36569	.132	-32.8343	453.6343
		Punjabi	234.50000	84.36569	.066	-8.7343	477.7343
	Sindhi	Balochi	24.00000	90.55759	1.000	-237.0862	285.0862
		Pashto	-701.33333*	84.36569	.000	-944.5677	-458.0990
		Saraiki	194.46667	84.36569	.198	-48.7677	437.7010
		Urdu	186.40000	90.55759	.314	-74.6862	447.4862
	Balochi	Punjabi	210.50000	90.55759	.190	-50.5862	471.5862

		Sindhi	-24.00000	90.55759	1.000	-285.0862	237.0862
		Pashto	-725.33333*	90.55759	.000	-986.4195	-464.2472
		Saraiki	170.46667	90.55759	.417	-90.6195	431.5528
		Urdu	911.73333*	84.36569	.000	668.4990	1154.9677
		Punjabi	935.83333*	84.36569	.000	692.5990	1179.0677
	Pashto	Sindhi	701.33333*	84.36569	.000	458.0990	944.5677
		Balochi	725.33333*	90.55759	.000	464.2472	986.4195
		Saraiki	895.80000*	84.36569	.000	652.5657	1139.0343
		Urdu	15.93333	84.36569	1.000	-227.3010	259.1677
		Punjabi	40.03333	84.36569	.997	-203.2010	283.2677
		Sindhi	-194.46667	84.36569	.198	-437.7010	48.7677
	Saraiki	Balochi	-170.46667	90.55759	.417	-431.5528	90.6195
		Pashto	-895.80000*	84.36569	.000	-1139.0343	-652.5657

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

U_U.F1

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Pashto	30	423.7000	
Balochi	23	424.7391	
Saraiki	30	444.8000	
Urdu	30	447.2000	
Punjabi	30	447.5333	
Sindhi	30		530.7000
Sig.		.822	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 28.552.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

U_U.F2

Tukey HSD

Group	N	Subset for alpha = 0.05	
		1	2
Punjabi	30	1096.5000	
Urdu	30	1120.6000	
Saraiki	30	1136.5333	
Balochi	23	1307.0000	
Sindhi	30	1331.0000	
Pashto	30		2032.3333
Sig.		.078	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 28.552.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.