Available online at www.jlls.org

JOURNAL OF LANGUAGE AND LINGUISTIC STUDIES

ISSN: 1305-578X

Journal of Language and Linguistic Studies, 17(Special Issue 1), 517-533; 2021

Phonemic variations in similar words of Turkish and Urdu language

Tania Ali Khan[a1]

[a]*Minhaj University, Pakistan*

APA Citation:

Khan, T.A. (2021). Phonemic variations in similar words of Turkish and Urdu language. *Journal of Language and Linguistic Studies, 17*(Special Issue 1), 517-533.

Submission Date:11/09/2020

Acceptance Date:.23/12/2020

Abstract

Urdu language is a member of Indo-European family tree and within the zone of Indo-Iranian branch, whereas Turkish language is a member of Altaic family tree. Both of these languages belong to different family trees, but these languages have many words in common. Urdu language has 41 consonant sounds and 11 vowel sounds, whereas Turkish language has 21 consonant sounds and 8 vowel sounds. Both of these languages don't have same number of phonemes. It is interesting to contemplate, how the speakers of both languages produce and perceive these common words in their languages. Therefore, a study is designed to explore the phonemic variations in similar words of Urdu and Turkish languages. In order to find out variations, a list of 75 words, which are common in Urdu and Turkish language was prepared in form of text. The data of this study was collected from Pak-Turk school of Lahore. The prepared words list was given to 10 Turkish speakers. These speakers were selected by using purposive sampling technique. The audios of the speakers were recorded and transcribed into standard Turkish IPA symbols. Levenshtein algorithm framework is used to draw a comparative analysis of Turkish phonemic transcriptions with standard Urdu phonemic transcriptions of the same words. With the help of Levenshtein algorithm, phonemic variations in similar words of both languages was measured. The intriguing result of the study will help in establishing the understanding about the production directions of the similar words used by the speakers of both languages.

Keywords: Urdu language; Turkish language; phonemes variations; Levenshtein algorithm

1. Introduction

Urdu language belong to Indo-European family tree and within the same family tree it belong to Indo-Iranian branch of languages. It has 193,238, 868 speakers in Pakistan and it is the official language of the country along with English language. Due to the fact of the neighbourhood relationship in these geographies, Urdu language has taken a lot of words from Turkish, Persian, Arabic, Hindi, Chinese and other languages. (Gracia & Yapici, 2014). As a matter of fact, evolution period of Urdu language was started with the contact of these neighbouring languages and out of these neighbouring languages the significant one's are Turkish, Persian Arabic and Hindi languages. (Khan, 2010). Indeed Urdu word itself is a Turkish language word, which mean is "an army". Urdu has a very rich phonetic inventory 13,

[1] Corresponding author.
E-mail address: umt_0002@hotmail.com

combination of Urdu letters and diacritics realizes 44 consonants, 28 non-aspirated & 16 aspirated. (Saleem et al., 2002). Furthermore, it has 22 stops, 8 fricative, 5 nasals and 6 liquids and glides. Moreover, it has 11 vowels short and long oral and nasal. (Raza, Hussain, Sarfraz, Ullah & Sarfraz, 2009).

Turkish language is a member of Altaic language family. It has 50 million speakers. It is the national language of Turkey. This language has taken words from Arabic and Persian. Turkish language is considered an orthographic language, a language that is written the way it is pronounced, or pronounced the way it is written (Yavuz & Balci, 2011). Turkish phonetic system uses 21 consonant sounds and 8 vowels. Turkish language also characterized its vowels as long vowels, which source came from the Arabic and Persian language (Ven Der Hulst & Ven De Weijer, 1991). In Turkish language majority of the syllables have CVC structure. And diphthongs do not exist in Turkish language. (Coldemil, 2018).

Although there are many difference between two languages like both belong to different language families, use different script systems, have different grammar structure. Moreover, Urdu language has nasality and timing and Turkish has nothing like that. (Younas, 2012). In contrast to this Turkish language has vowel harmony and Agglutination and Urdu has nothing like that. (Coldemil, 2018). But these two languages have some elements in common, like Turkish and Urdu languages commonly share two categories of words. First category is comprised of words, which Urdu language has directly borrowed from Turkish language. And these words are round 2608. According to Dr Syed Mohammad Anwer, from these 2608 words only 24 words are purely Turkish Words. The second category is comprised of words, which came in Urdu and Turkish languages from Arabic and Persian languages. (Younas, 2012).

In addition to this, both languages share the same word order “SOV” and all six combinations of SOV is possible in these languages, therefore these languages are known as semi-free order languages. (Coldemil, 2018). This is a rather growing idea in the field of phonetics and phonology to explore the phonemic differences across different dialects of the one language and also explore the phonetic settings. But recently, it seemed as researchers are interested to explore the phonemic variations across languages.

Recently, it has been noticed that researchers are interested to explore the phonemic variations across languages. In connection of Urdu and Turkish languages many studies have discussed the case of similar words in both languages and acknowledged the difference in the production. But the phonemic variations in the similar words have not been discussed yet. Khattak wrote in his book "Urdu aur Turki ke Mushtarak Alfaz" in 1987 (Similar words of Urdu and Turkish) that, there are 2608 common words which are spoken by Turkish and Urdu speakers; however there is a bit difference in their production. (Manwar, 2011). Therefore, the current study focuses on the distinctions in the production of the similar words that exist in Turkish and Urdu languages. Hence, the researcher used Levenshtein algorithm to find out the phonemic variations between Urdu and Turkish language’s vowels and consonants sounds in the production of similar words.

1.1. Literature review

This is an open secret now that all the languages around the world use different phonemes and phonologies to distinguish themselves. But this realization was not as easy as it seems. It came with tenacious question like Does the earlier human demographic history leave human race with similar signature phonemes? It took the ages to find out the answer of this question that worldwide total number of phonemic inventories are 2,082. (Creanza, Ruhlen, Pemberton, Rosenberg, Feldman & Ramachandran, 2015). But this figure of phonemic inventories seems to increase now. In 2019 a cross linguistic phonological inventory database released “PHOIBLE”. In this database the total number of phonemic inventories is 3020 that contain 3183 segment types found in 2186 distinct languages. (Moran

& McCloy, 2019). All of this complicated facts make it meaningful to discuss the concept of Phoneme. "A phoneme is the unit of sound that distinguish one word from another in a particular language". (Roach, 2009). Another definition "The smallest unit of sound that makes a difference in communication." (Yopp & Yopp, 2000)

This smallest unit of sounds are further categorized into vowels and consonants. Roach stated that consonants are the phonemes, articulated with complete and partial closure of the vocal tract. And vowels are the sounds which are articulated with any stricture in the vocal cavity. (Roach, 2009). This smallest unit of sound incorporates in the construction of syllables and words. These phonemes are manipulated by stretching, blending segmenting, isolating substituting and deleting in different languages around world. (Vogt & Shearer, 2011). This stretching, blending segmenting, isolating, substituting and deleting became the cause of phonemic variations.

In some languages phonemic variations are significantly visible and in some languages it is not. It is due to the fact of phonemic distance between the populations that is correlated with geographic distance and which leads to languages distance. (Creanza et al., 2015) Geographically close languages show less variations in phonemes. However; as the distance increased phonemic difference also increased between languages. Additionally, it has been seen that geographically isolated languages have less or no phonemic diversity by reason of no or less contact with other languages. Therefore, the phonemes have no fear to drift. On the contrary, languages with more neighbours seemed to face the fear of phoneme change substantially. (Mennen, Leeuw, Scobbie & Schaeffler, 2010).

1.2. Inventories of Urdu language

In case of Urdu and Turkish languages, these belong to two different geographical locations. The evident distance between the two languages indicates that the phonemic variations between the languages will be of immense attention. Phonetically, Urdu is a rich language with a large inventory of consonants and numerous long nasal, long non-nasal. (Raza et al., 2009) Hussain stated that in Urdu language, vowels use articulation, nasality, duration very distinctively. Short vowels are represented by diacritics, for instance zer, zabar, pesh. Most of the words end with long vowels sounds in Urdu language. (Hussain, 2004). Urdu has 41 consonants in total, including stops and affricates, fricatives, nasals, and liquids/glides. The stops and nasals are articulated at five different places, being classified as labial, dental, retroflex, palatal and velar. The palatal stops are, in fact, affricates. Every series of stops includes voiceless and voiced consonants, unaspirated and aspirated, this four-way contrast being unique to Indo-Aryan among Indo-European languages. (Ejaz, 2013).

		Labial	Dental	Retroflex	Palatal	Velar	Uvular	Glottal
Stop	Voiceless	p pʰ	t tʰ	ʈ ʈʰ		k kʰ	q	ʔ
	Voiced	b bʰ	d dʰ	ɖ ɖʰ		g gʰ		
Affricate	Voiceless				ʧ ʧʰ			
	Voiced				ʤ ʤʰ			
Fricative	Voiceless	f	s		ʃ	x		h
	Voiced		z		ʒ	ɣ		
Nasal		m	n	ɳ	ɲ	ŋ		
Liquid			l r	ɽ ɽʰ				
Glide		ʋ			j			

Figure 1. IPA table of Urdu Language Consonants

Urdu language has 11 vowel systems composed of three lax and eight tense vowels. Lax vowels (ɪ, ʊ, ə) are phonetically short and tense vowels (i, e, ɛ, u, o, ɔ, ɑ) are phonetically long. [ɪ] is slightly lower

and more centralized than [i], [ʊ] is slightly lower and more centralized than [u]. All have nasal forms. Oral and nasal vowels are contrastive. (Raiz, 2001).

	Front	Central	Back
High	i ɪ		ʊ u
High-mid	e		o
Mid		ə	ɔ
Low-mid	ɛ		
Low			ɑ

Figure 2. IPA table of Urdu language vowels

1.3. Inventories of Turkish language

Turkish language has eight systematic vowels and 21 consonants. Lip rounding is the most distinctive feature of Turkish vowels. (Levi, 2001). Turkish language has nasals, stops, affricates, fricatives, approximants, flap and laterals. Like other languages Turkish language also uses the letter <l/> to represent a lateral sound. Unlike Urdu language, Turkish language clearly differentiates between the two different variants of this sound, i.e., the velarized lateral and non-velarized lateral. And these two different lateral phonemes are greatly affected by their surrounding vowels (Börtlü, 2020). Moreover, in Turkish language, if a word ends with a consonant, then it will be a voiceless consonant. Another unique feature of Turkish language consonants is that they do not use aspirated sounds. (Clements & Sezer, 1982).

Consonant phonemes of Standard Turkish								
Place of articulation Manner of articulation	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Nasal	m		n				
Stops	p b		t d			k g	
Affricate				t͡ʃ d͡ʒ			
Fricative	f		s z	ʃ ʒ			h
Lateral		ɫ̪	l				
Approximant		v			j		
Flap			ɾ				

Figure 3. IPA table of Turkish language consonants

Turkish vowels use such features as labial for rounded vowels, coronal for front vowels and dorsal for back vowels. It has also been found out that all of these features are vocalic in Turkish language. (Hunter, 2013). In Turkish language vowels have harmony that causes vowels in most words to be either front or back of the mouth and either rounded or unrounded. Turkish language does not use diphthongs. When two vowels are adjacent in the syllable of a word, each vowel retains its individual sound. (Hargus, 2011).

1.4. Phonetics differences

There are some sounds in Urdu and Turkish language that have different symbols for instance, the sounds (cha) or (ja) are represented with different symbols in both languages. Turkish consonants are

divided into voiced and voiceless consonants. But the voiceless consonants are pronounced with little aspirated sound which cannot describe distinctively. (Levi, 2001). Unlike any other language, Urdu language uses aspirated sounds distinctively and assigns a unique character to them, not only in script form but also in IPA symbols. (Ejaz, 2013). Urdu language differentiates between the bilabial and labio-dental sounds. Moreover, it takes three nasal sounds, m,n,ŋ. (Hussain, 2004). A study conducted on nasal aspirants in Urdu language revealed that [n^h] and [m^h] sounds do not occur in words at initial and final positions. (Shah, 2002).

This study is a cross linguistics study and two languages Urdu and Turkish are under observation of the current paper. Cross-linguistics similarities are reflected by the speech sound systems of languages all over the world. Therefore, the consonants and vowels inventory size distribution and their preferred attachments is the ultimate reason behind the emergence of such a study. (Choudhary et al., 2006).

The researcher noticed these interesting and unique pronunciation patterns in speech of Turkish people, who are inhabitants of Pakistan. This element led this study to figure out and analysed the changes that exist in the articulation of the similar words in Urdu and Turkish languages made by the Turkish speakers. Therefore, a word to word mapping approach was used for this study but a word to word mapping analysis get wrong transliteration of the words because words have different spellings in both languages. Solution to this problem is to maintain a list of words that differ in spelling in both languages. (Jawaid & Ahmed, 2009)

1.5. Levenshtein Distance

It is a method to measure the distance between the two strings. This distance is measured in term insertion, deletion and substitution of one or more strings. The more the number of Levenshtein distance, the more the differences. (Chohan, Habib & Hasan, 2020) It is a favoured method for comparing the whole word of language to the other word of the language for the intention of computing the distance. (Kessler, 2005). In 1965, a Russian scientist, Vladimir Levenshtein invented this method. In this method cost is assigned to a pair that do match up; this is known as substitution. Cost is assigned to each string that does not match up; this is known as deletion or insertion. (Luce & Pisoni, 1998). Following example explains this algorithm, as "kitten" and "sitting" have a difference of three strings.

1. Kitten → sitten (substitution of "s" for "k")
2. sitten → sittin (substitution of "i" for "e")
3. sittin → sitting (insertion of "g" at the end). (Chohan et al., 2020)

1.6. Research questions

As the study aims to find consonants and vowels variations in similar words of Turkish and Urdu languages, the following are the research questions

1. What are the consonantal variations in the production difference of the similar words of Urdu and Turkish language?
2. What are the vowels variations in production of the similar words of Urdu and Turkish language?

2. Method

The mixed method is used to analyse the phonemic variations in Urdu and Turkish languages. The existing phonetic inventory of Urdu and Turkish languages are in the form of words to find out phonemic variations. Moreover, it is a cross-sectional study as the data was collected in a small period of time.

2.1. Research Instrument

The researcher prepared the list of 75 words, which are common in both languages. The selection of these words is totally according to the IPA charts of Urdu and Turkish languages.

2.2. Target Population

The target population of this study was all the participants of Pak-turk institute of Lahore.

2.3. Sample size & Sampling technique

These participants of this study were selected with non-random purposive sampling technique. The sample size of the participants were 10 female Turkish students.

2.4. Data collection

The data of this study was collected from participant of Pak-Turk institute of Lahore. It was collected in form of audios. The participants were given the list of the words and the audios were recorded. It was made sure that these words are known to the sampled population.

2.5. Framework/Analysis

This study used Levenshtein algorithm model as the framework of this research. This algorithm gives a decisive picture of phonetics similarities and differences in two different linguistic systems. The selected words of the languages are transcribed into their standard IPA symbols and then the comparison is drawn. (Maldonado García & Borges de Souza, 2014; Heeringa, 2004; Sanders & Chin, 2009). The phonemes of one word are mapped with the other and differences are measured in terms of substitution, deletion and insertion processes.

3. Results

Levenshtein Algorithm was used for the data analysis. According to this algorithm the number of distances among the consonants and vowels sounds of Turkish and Urdu language was calculated. Moreover, the number of the similar consonants and vowels sounds between the two languages was also calculated. The selected words were transcribed into IPA symbol of Urdu and Turkish languages. To confirm the transcription of the Turkish words, recorded voices of the speakers were also transcribed. And then comparison was drawn between the transcriptions of both languages by using Levenshtein Algorithm. Vowels and consonants sounds of both language were mapped. The analysis gives the distance in the term of numbers between two sounds. Value 1 was allotted to the different sounds. And Value 0 was used for the similar sounds. This distance and similarities were calculated between Turkish and Urdu language in the term of their phonemes.

3.1. Same and different consonant sounds in both languages

Turkish and Urdu languages are different languages, which belong to different family trees. There are chances of the maximum phonemic differences between the two languages. The consonant sounds which are same in both languages have zero distance. But the sound which does not exist in either of the language have 1 distance. A substitute sound is used in place of that sound or that particular sound is deleted while pouncing the word. It can be seen in Table 1.

Table 1. Same & different consonant sounds in Turkish and Urdu language

Sr#	Consonants Sounds in Turkish	Consonants Sounds in Urdu	Levenstein Distance
1	/m/	/m/	0
2	/n/	/n/	0
3		/ŋ/	1
4	/p/	/p/	0
5		/pʰ/	1
6	/b/	/b/	0
7		/bʱ/	1
8		/ʈ/	1
9		/ʈʰ/	1
10		/ɖ/	1
11		/ɖʱ/	1
12	/t/	/t/	0
13	/d/	/d/	0
14		/tʰ/	1
15		/dʰ/	1
16	/k/	/k/	0
17	/g/	/g/	0
18		/kʰ/	1
19		/gʰ/	1
20		/q/	1
21	/f/	/f/	0
22	/v/	/v/	0
23	/s/	/s/	0
24	/z/	/z/	0
25	/ʃ/	/ʃ/	0
26	/ʒ/	/ʒ/	0
27		/X/	1
28		/ɣ/	1
29	/h/	/h/	0
30	/tʃ/	/tʃ/	0
31	/dʒ/	/dʒ/	0

32		/tʃʰ/	1
33		/dʒʱ/	1
34		/rʰ/	1
35		/r/	1
36	/ɽ/	/ɽ/	0
37	/j/	/j/	0
38	/l/	/l/	0
39		/ɲ/	1
40		/ɳ/	1
41		/ʔ/	1
42	ɬ		1
43	(c)		1
44	(ɟ)		1

Total similar consonants sounds in both languages are 18 and different sounds in both languages are 24. Table 2 provides the percentage in the number of similarities in sounds and distances in the consonant sounds.

Table 2. Total number of sounds with percentage

Consonant sounds	Turkish & Urdu language
Total similar consonants sounds in Turkish and Urdu language	45.45 %
Total different consonants sounds in Urdu and Turkish language	54.54%

This analysis revealed that there is 45.45% phonemic similarity in Turkish and Urdu language. There is a 54.54% ratio in connection with their distance. This similarity and difference explains Turkish and Urdu language have less similarity and more difference between them.

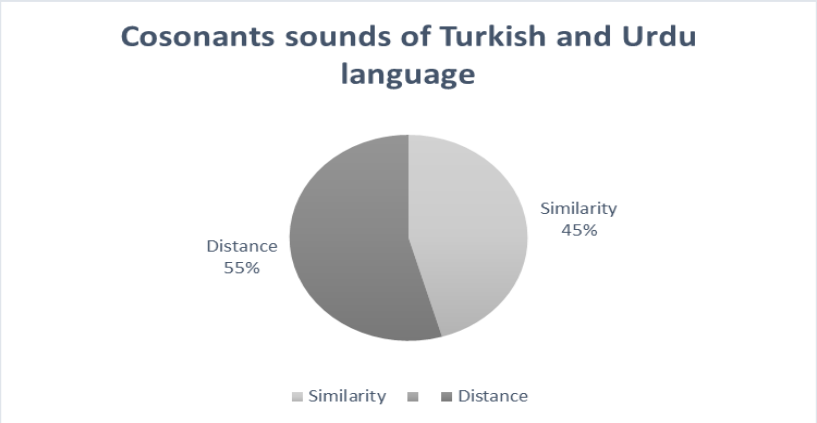

Figure 1. Similarity and difference index

The above figure explains difference and similarity of index in the Turkish and Urdu languages.

3.2. Same and different vowels sounds in both languages

The vowel sounds which are same in both languages have zero distance. But the vowel sound which does not exist in either of the language have 1 distance. A substitute sound is used in place of that sound or that particular sound is deleted while pouncing the word. It can be seen in Table 3.

Table 3. Same and different vowel sounds in Turkish and Urdu language

Sr#	Vowels Sounds in Turkish	Vowels Sounds in Urdu	Levenstein Distance
1	/i/	/i/	0
2	/I/	/I/	0
3	/e/	/e/	0
4		/ɛ/	1
5		/æ/	1
6	/a/	/a/	0
7	/u/	/u/	0
8	/o/	/o/	0
9		/ə/	1
10		/ɔ/	1
11		/ʊ/	1
12	/ü/		1
13	/ŏ/		1

Total similar vowel sounds in both languages are 6 and different sounds in both languages are 7. Table 4 provides the percentage in the number of similarities in sounds and distances in the vowel sounds.

Table 4. Total number of vowel sounds with percentage

Vowels sounds	Turkish & Urdu language
Total similar Vowels sounds	46.15%
Total different vowels sounds	53.85%

This analysis revealed that there is 46.15% phonemic similarity in Turkish and Urdu language. There is a 53.85% ratio in connection with their distance. This similarity and difference explains Turkish and Urdu language have less similarity and more difference between their vowel sounds.

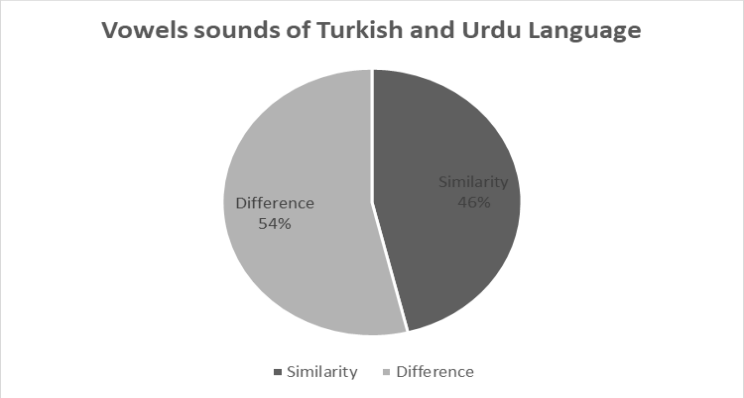

Figure 2. Similarity and difference index

The above figure explains difference and similarity of index in the Turkish and Urdu languages vowel sounds.

4. Discussion

4.1. Total number of distance and similar phonemics in Turkish and Urdu language

During the analysis, It was observed that Urdu language consonants and vowels which exist in Turkish language, participants felt no problem in uttering those consonantal and vowels sounds, but the consonants and vowels of Urdu language which do not exist in Turkish language, Turkish speakers switched them with their own articulatory equivalent consonants and vowels. Moreover, it was also noticed that Turkish students can easily pronounce all vowels in the Urdu language. Table 5 shows the total number of distance and similar consonants and vowels sounds in Turkish and Urdu language.

Table 5. Total number of distance and similar sounds

Consonants and vowels	Levenshetein distance	Number of sounds
Distance consonant sounds	1	24
Similar consonant sounds	0	18
Distance vowel sounds	1	7
Similar vowel sounds	0	6

It helped in determining the tables of sounds changing rules used by Turkish speakers. These are as follow.

- Consonants sound changing rules used by Turkish speakers in similar words.
- Vowels sounds changing rules used by Turkish speakers in similar words.

4.2. Case of Consonant sounds in Turkish and Urdu language

It was observed that if Voiced Bilabial /b/ sound is coming at the end of word, it is switched into voiceless bilabial /p/ by the Turkish speakers. And Voiced Dental /d̪/ at the end of the word changed into voiceless dental /t̪/. As for Nasal /ɳ/ does not exists in Turkish sound system therefore, it changed into nasal /n/. Glottal /h/ at the end of word changed into vowel sound. Sound /tʃh/ at initial position in word changed into unaspirated /tʃ/. Velar fricative /ɣ/, /X/ do not exists in Turkish language therefore,

speakers changed these sounds into velar voiced stops that is /g/ Velar fricative /ɣ/ at initial position in a word changed into velar /g/.

Table 6. Case of consonants in similar words

Urdu consonants sounds switched in Turkish language	
Urdu Consonantal Sounds	Turkish Consonantal Sounds
Voiced Bilabial /b/ sound at the end of word	Changed into voiceless bilabial /p/
Voiced Dental /d̪/ at the end of the word	Changed into voiceless dental /t̪/
Nasal /ŋ/	Changed into nasal /n/
Glottal /h/ at the end of end	Changed into vowel sound
/ʧʰ/ at initial position in word	Changed into unaspirated /ʧ/
Velar fricative voiced /ɣ/ at initial position in word	Changed into velar / g/

Data shows that there are some of the variations in the pronunciation of consonants at the end of the words. Turkish language does not license voiced consonants at the end of the words. Turkish language alphabets have the ğ sound but it is always silent. (Demircioglu, 2013) When bilabial /p/, dental /t̪/, /k/ and palatal /ʧ/ comes at the beginning of the words they pronounce with aspirated sounds like English language. Turkish language only possess aspirated version of these voiceless sounds. (Kallestinova, 2009) Therefore, the speakers feel no problem in uttering aspirated sounds in initial position of word. But they feel problem in uttering unaspirated version of these voiceless sounds. In addition uvular consonant do not exist in Turkish language there are only two nasal consonants i.e. /m/, /n/. If the first syllable end with diphthong the second syllable start with alveolar /d/ consonant.

4.3. Case of Vowel sounds in Turkish and Urdu language

While analysing vowel sounds, it is observed that the speakers converted short vowel /ʊ/ into their rounded vowels /Ü/, /Ö/, /i/. Back low vowel /a/ came after voiced consonant changed into short/ə/. Back low vowel /a/ came after voiceless consonant at the end of word changed into short /e/. Short vowel /ə/ changed into unrounded from of vowels /e/, /ɪ/, /i/, and Front vowel in the center of word /ɜ/ changed into back low /a/.

Table 7. Case of vowels in similar words

Urdu vowels	Turkish vowels
Short vowel /ʊ/	Changed into their rounded vowels /Ü/, /Ö/, /i/
Back low vowel /a/ came after voiced consonant	Changed into short/ə/
Back low vowel /a/ came after voiceless consonant at the end of word	Changed into short /e/
Back low vowel /æ/	Changed into /a/
Short vowel /ə/	Changed into unrounded from of vowels /e/, /ɪ/,/i/,
Front vowel in the center of word/ɜ/	Changed into back low /a/

In case of vowels, it was noticed that lip rounding in vowels pronunciation is the distinctive feature of Turkish sound system. (Arik, 2015) The data shows that Turkish speakers often use lip rounded vowels when they are required to produce Urdu long vowels. At the end of the words, if there is a consonant

before the vowel then the vowel will be an unrounded vowel. But if the word is ending with a vowel in Turkish context of common words then the vowel will be rounded and short one. Whereas, Urdu language has the same quality but, it does not license the short vowel at the end of the word. Turkish language speakers pronounce words quickly than the speakers of Urdu language.

5. Conclusions

On the basis of the analysis, it is concluded that although both languages share limited number of vocabulary with exactly the same meaning but there is clear difference in their pronunciation. The Urdu language speakers used those consonants and vowel sounds which are licensed by Urdu language. And Turkish speakers used those sounds which are allowed in their phonetic settings. Only 18 consonant sounds are common in both languages and 24 consonant sounds are different in two languages. In terms of vowels, 6 vowel sounds are same in both languages however, 7 vowel sounds are different. There are some consonants and vowels which exist on the same place of articulation in both languages, however, no aspirated forms are found of some consonants in Turkish language.

Turkish language does not used aspirated sounds but speakers seem to face no problem in uttering the sounds because some Turkish consonants only takes aspirated forms of English words, like /p^h/, /t^h/, /k^h/. The Turkish speakers pronounce these sounds aspirated only when they come at the initial position of the words. In addition, diphthongs do not exist in Turkish language but when it comes to pronouncing the Urdu words, the Turkish speakers do not find any difficulty regarding articulation of these diphthongs.

It has also been examined in some cases that the Turkish speakers end a word with the articulatory equal but different sound in contrast of Urdu language. Therefore, it is found out that both of the languages have same articulation equivalence. It is observed that Turkish speakers as L2 learners of Urdu language face some problems at the initial level of learning Urdu language, but the learners who have more exposure of Urdu language and frequently communicate with native people on day to day bases overcome on these linguistics barriers.

6. Ethics Committee Approval

The author(s) confirm(s) that the study does not need ethics committee approval according to the research integrity rules in their country (Date of Confirmation: January 14, 2021).

References

Arik, E. (2015). An experimental study of Turkish vowel harmony. *Poznan Studies in Contemporary Linguistics, 51*(3), 359-374.

Börtlü, G. (2020). The second formants of the laterals in Turkish. *Journal of Language and Linguistics Studies, 16*(2), 510-520.

Chohan, M. N., Habib, M. A., & Hasan, W. (2020). Phonemic comparison of Majhi and Shahpuri-dialects of Punjabi. *Hamdard Islamicus, 43*(2), 104-114.

Choudhury, M., Mukherjee, A., Basu, A., & Ganguly, N. (2006). Analysis and synthesis of the distribution of consonants over languages: A complex network approach. *Coling: Computational Linguistic, 6*(1), 128-135.

Clements, G. N., & Sezer, E. (1982). *Vowel and consonant disharmony in Turkish* [Doctoral dissertation].http://linguistics.ucla.edu/people/hayes/251VowelHarmony/readings/ClementsSezer1982TurkishVowelHarmony.pdf

Coldemil, S. (2018). Similar words between Turkish & Hindi (Urdu) languages in terms of the open and closed class morphemes. *Sosyal Bilimler Enstitüsü Dergisi, 8*(1), 190-199.

Creanza, N., Ruhlen, M., Pemberton, T., Rosenberg, N., Feldman, M., & Ramachandran, S. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *Proceedings of the national Academy of sciences of United State of America, 112*(5), 1265-1272.

Demircioglu, M. D. (2013). The pronunciation problems for Turkish learners in articulating of the diphthongs in English. *Procedia - Social and Behavioral Sciences*, *106*(1), 2983-2992.

Ejaz, M. (2013). Phonological behavior of multiple aspirated consonants in Urdu. *Center of Language Engineering, 1*, 54-61.

Hargus, P. (2011). Turkish Vowel harmony: Under specification, iteration, multiple rules [PowerPoint slides]. Retrieved from http://webcache.googleusercontent.com/search?q=cache:Pooc_gQvsO4J:courses.washington.edu/lingclas/451/Turkish_vowel_h

Heeringa, W. J. (2004). *Measuring dialect pronunciation differences using Levenshtein distance*. [Unpublished doctoral dissertation]. University of Groningen.

Hunter, H. S. (2013). *A study of the Turkish vowel harmony and the power of language* [Unpublished master's thesis]. Kent State University Honors College.

Hussain, S. (2004). Letter to sound rules for Urdu text to speech system. *Center for Research in Urdu Language Processing, 1*, 1-6.

Jawaid, B., & Ahmed, T. (2009). Hindi to Urdu conversion: Beyond simple transliteration. *Proceedings of the Conference on Language & Technology, 1*(1), 24–31.

Kallestinova, E. (2009). Voice and aspiration of stops in Turkish. *Folia Linguistica, 38*(1-2), 117-129.

Kessler, B. (2005). Phonetic Comparison Algorithms. *Transactions of the Philological Society, 103*(2), 243-260.

Khan, U. (2010). Of Turkish loanwords in Urdu and Dr. Sabir. The dawn, p. 2. https://www.dawn.com/news/582032/of-turkish-loanwords-in-urdu-and-dr-sabir-2

Levi, S. V. (2001). Glides, laterals, and Turkish vowel harmony. *The Chicago Linguistic Society, 37*(1), 379-393.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *National Institute of Public Health Public Access, 19*(1), 1-36.

Maldonado García, M., & Borges De Souza, A. (2014). Lexical similarity level between English and Portuguese. *Grassroots, 49*(1), 203-218.

Maldonado, M. I., Garcia, & Yapici, M. (2014). Common vocabulary in Urdu and Turkish language: A case of historical onomasiology. *Journal of Pakistan Vision, 15*(1), 1-33.

Manwar, S. (2011, November 13). Language: Urdu and the borrowed words. The Dawn, p. 4. https://www.dawn.com/news/672945/language-urdu-and-the-borrowed-words

Mennen, I., Scobbie, J., Schaeffler, F. & Leeuw, E. (2010). Measuring language-specific phonetic settings. *Second language Research, 26*(1), 13-41.

Moran, S., & McCloy, D. (2019). Phoible 2.0. Retrieved from https://phoible.org/

Raiz, K. (2001). Synthesis of oral and nasal vowels of Urdu. *Center of Language Engineering*, *1*, 94-101.

Raza, A. A., Hussain, S., Sarfraz, H., Ullah, I., & Sarfraz, Z. (2009). Design and development of phonetically rich Urdu speech corpus. *Urdu Speech Corpus, 9*(1), 38-43.

Roach, P. (2009). *Phonetics and Phonology* (4th ed.). Cambridge, U.K: Cambridge University Press.

Saleem, A. M., Kabir, H. A. S. A. N., Riaz, M. K., Rafique, M. M., Khalid, N. A. U. M. A. N., & Shahid, S. R. (2002). Urdu consonantal and vocalic sounds. *CRULP Annual Student Report*.

Sanders, N. C., & Chin, S. B. (2009). Phonological distance measures*. *Journal of Quantitative Linguistics, 16*(1), 96–114.

Shah, S. M. (2002). Urdu nasal consonants and their phonological behavior. *Center of Language and Engineering, 1*, 133-140.

Van der Hulst, H., & Van De Weijer, J. (1991). Topics in Turkish phonology. *Turkish Linguistics Today, 1*, 11-59.

Vogot, M.E. & Shearer, B. (2011). *Reading specialists and literacy coaches in the real world* (3rd Ed.). Boston, MA: Allyn and Bacon.

Yavuz, H., & Balcı, A. (2011). *Turkish phonology and morphology* (1st ed.). Eskişehir: Anadolu University.

Yopp, R., & Yopp, H. (2000). Supporting phonemic awareness development in the classroom. *The Reading Teacher, 54*(2), 130-143.

Younas, A. (2012). Case of shared vocabulary of Turkish and Urdu languages (similar words in Turkish and Urdu languages. *SSRN Journal, 10*, 1-22.

Appendix A.

Table of Edit distance between the similar words of Turkish and Urdu language

No	Words	Meanings	Turkish Transcription	Urdu Transcription	Levenshtein distance	Phonemic options in Turkish	Phonemic options in Urdu
1	Barud	Gunpowder	/b**ɑ**ruṯ/	/b**a**ruḏ/	2	(2) Substitution	(2) Substitution
2	Bulbul	Nightingale	/ b**ü**lb**ü**l/	/b**ʊ**lb**ʊ**l/	2	(2) Substitution	(2) Substitution
3	Badam	Almond	/baḍəm/	/baḍ**a**m/	1	Substitution	Substitution
4	Bazoo	Arm	/**p**azʊ/	/**b**az**u**/	2	(2) Substitution	(2) Substitution
5	pehelvan	Wrestler	/p**e**hl**ɪ**van/	/p**ə**hələvan/	2	(1) Substitution & (1) insertion	(1) Substitution & (1) insertion
6	Pulow	Rice dish	/p**i**la**v**/	/pəla**o**/	2	(2) Substitution	(2) Substitution
7	Musfir	Traveler	/m**i**safır/	/m**ʊ**safır/	1	Substitution	Substitution
8	Musibat	Trouble	/m**ü**sibəṭ/	/m**ʊ**sibəṭ/	1	Substitution	Substitution
9	Muhabbat	Love	/m**ü**həbəṭ/	/m**ʊ**h**a**bəṭ/	2	(2) Substitution	(2) Substitution
10	Meidan	Ground	/meiḍan/	/meḍan/	1	Insertion	deletion

11	Fakir	Beggar	/fəkir/	/fəkir/	0	Nil	Nil
12	Fazool	Objectless	/füzül/	/fəzul/	2	Substitution	Substitution
13	Fidha	Benefit	/faiḍa/	/faeḍa/	1	Substitution	Substitution
14	Watan	Country	/vəṭən/	/vəṭən/	0	Nil	Nil
15	Valdha	Mother	/valıḍæ/	/valḍah/	3	(1) Insertion, (1) substitution and (1) deletion	(1) Deletion, (1) substitution and (1) insertion
16	Tasdique	Confirm	/ṭəsḍik/	/ṭəsḍik/	0	Nil	Nil
17	Tava	Pane	/ṭəva/	/ṭəva/	0	Nil	Nil
18	Tazha	Fresh	/ṭaze/	/ṭaza/	1	Substitution	Substitution
19	Toop	Cannon	/ṭöp/	/ṭop/	1	Substitution	Substitution
20	Tabancha	Gun	/ṭəbanʤa/	/ṭəbənʧa/	2	(2) Substitution	(2) Substitution
21	Dafa	Turn	/ḍəfa/	/ḍəfa/	0	Nil	Nil
22	Dard	Pain	/ḍərṭ/	/ḍarḍ/	2	(2) Substitution	(2) Substitution
23	Diwar	Wall	/ ḍüvar/	/ḍivar/	1	Substitution	Substitution
24	Dukan	Shop	/ḍükan/	/ḍukan/	1	Substitution	Substitution
25	Duniya	World	/ḍünja/	/ḍʊnıja/	2	(1) Substitution & (1)deletion	(1) Substitution & (1) insertion
26	Doorbin	Telescope	/ḍürbün/	/ ḍurbin/	2	(2) Substitution	(2) Substitution
27	Dushman	Enemy	/ḍüʃmən/	/ ḍʊʃmən/	1	Substitution	Substitution
28	Anar	Pomegranate	/nar/	/anar/	1	Deletion	Insertion
29	Nafrat	Hatred	/nefrət/	/nəfrət/	1	Substitution	Substitution
30	Nijat	Immunity	/nıjəṭ/	/nijəṯ/	2	(2) Substitution	(2) Substitution
31	Saf	Clear	/saf/	/saf/	0	Nil	Nil
32	Sabzee	Vegetables	/sebze/	/səbzi/	2	(2) Substitution	(2) Substitution
33	Sabun	Soap	/sabon/	/sabʊn/	1	Substitution	Substitution
34	Zanjeer	Chain	/zəndʒir/	/zəndʒir/	0	Nil	Nil
35	Zamin	Earth	/zəmin/	/zəmin/	0	Nil	Nil
36	Zarab	Beat	/zərb/	/zərb/	0	Nil	Nil
37	Ruh	Soul	/ruh/	/rʊh/	1	Substitution	Substitution
38	Rahber	Guide	/ræhbər/	/ræhbar/	1	Substitution	Substitution
39	Rahna	To live	/ræhna/	/ræhna/	0	Nil	Nil
40	Chat	Roof	/ʧət/	/ʧʰəṯ/	2	(2) Substitution	(2) Substitution
41	Chakoo	Knife	/ʧəkŭ/	/ʧa.ku/	2	(2) Substitution	(2) Substitution
42	Chai	Tea	/ʧɑ.ɪ/	/ʧae/	2	(2) Substitution	(2) Substitution
43	Jawab	Answer	/ʤəvab/	/ʤʊab/1	1	Insertion	Deletion
44	Janam	Birth	/ʤɑnım/	/ʤənəm/	2	(2) Substitution	(2) Substitution
45	Sharab	Vain	/ʃərəp/	/ʃərɑb/	2	(2) Substitution	(2) Substitution
46	Shakar	Sugar	/ʃekər/	/ʃəkər/	1	Substitution	Substitution
47	Shakayat	Complain	/ʃıkɑjəṯ/	/ʃıkɑjəṯ/	0	Nil	Nil
48	Sheesha	Mirror	/ʃıʃa/	/ʃiʃɑ/	1	Substitution	Substitution
49	Yani	For instance	/jɑne/	/jɑni/	1	Substitution	Substitution
50	Yar	Beloved	/jɑr/	/jɑr/	0	Nil	Nil

51	Kitab	Book	/kɪt̪əp/	/kɪt̪ɑb/	2	(2) Substitution	(2) Substitution
52	Kofta	Meat dish	/kofṯə/	/kofṯɑ/	1	Substitution	Substitution
53	Keemah	Mince	/kima/	/kimɑ/	1	Substitution	Substitution
54	Katail	Killer	/kat̪ɪl/	/kɑt̪ɪl/	1	Substitution	Substitution
55	Gham	Sorrow	/gəm/	/ɣəm/	1	Substitution	Substitution
56	Garoor	Proud	/gűrʊr/	/ɣərur/	3	(3) Substitution	(3) Substitution
57	Gunah	Sin	/gűnəh/	/gʊnəh/	1	Substitution	Substitution
58	Helva	Sweet dish	/hælva/	/həlva/	1	Substitution	Substitution
59	Hesab	Calculation	/hisəp/	/hisab/	2	Substitution	Substitution
60	Hafta	Week	/həfṯɑ/	/həft̪əh/	2	Substitution & deletion	Substitution & insertion
61	Shatan	Devil	/ʃəjt̪ɑn/	/ʃɛt̪ɑn/	2	(1) Substitution & (1) insertion	(1) Substitution & (1) deletion
62	Lalh	Flower name	/lɑle/	/laləh/	3	(2) Substitution & (1) deletion	(2) Substitution & (1) insertion
63	Kalam	Pen	/kəlæm/	/kələm/	1	Substitution	Substitution
64	Tameez	Manner	/ẗ̪ʰemɪz/	/t̪əmiz/	3	(3) Substitution	(3) Substitution
65	Rahber	Guide	/ræhbər/	/rɛhbər/	1	Substitution	Substitution
66	Noor	Noor	/nʊr/	/nur/	1	Substitution	Substitution
67	Khadija	Khatija	/həṯidʒe/	/Xəḏidʒəh/	4	(3) Substitution & (1) deletion	(3) Substitution & (1) insertion
68	Fatima	Fatima	/fəṯma/	/faṯɪməh/	4	(2) Substitution & (2) deletion	(2) Substitution & (2) insertion
69	Hussain	Hussain	/hʊsɛɪn/	/hʊsæn/	1	Substitution	Substitution
70	Umair	Umar	/ömer/	/ʊmər/	2	(2)Substitution	(2) Substitution
71	Amir	Amir	/emɪr/	/ɑmɪr/	1	Substitution	Substitution
72	Arsalan	Arsalan	/əslən/	/ərsəlɑn/	3	(2) Deletion & (1) substitution	(2) Insertion & (1) deletion
73	Ahmad	Ahmad	/ɑhmeṯ/	/æhməḏ/	3	(3) Substitution	(3) Substitution
74	Firdous	Firdous	/fɪrḏ̪ əvs/	/fɪrḏos/	3	(2) Substitution & (1) insertion	(2) Substitution & (1) deletion
75	Aysha	Aysha	/ajʃe/	/ɑiʃəh/	3	(2) Substitution & (1) deletion	(2) Substitution & (1)insertion

Türkçe ve Urdu dilinin benzer kelimelerindeki fonemik varyasyonlar

Özet

Urdu dili, Hint-Avrupa soy ağacının bir üyesidir ve Hint-İran kolu bölgesinde yer alırken, Türk dili Altay soy ağacının bir üyesidir. Bu dillerin her ikisi de farklı aile ağaçlarına aittir, ancak bu dillerin birçok ortak kelimesi vardır. Urdu dilinde 41 ünsüz ve 11 sesli harf bulunurken, Türk dilinde 21 ünsüz ve 8 sesli harf bulunmaktadır. Bu dillerin ikisi de aynı sayıda ses birimine sahip değildir. Her iki dili konuşanların bu ortak kelimeleri kendi dillerinde nasıl ürettiklerini ve algıladıklarını düşünmek ilginçtir. Bu nedenle, Urduca ve Türk dillerinin benzer kelimelerindeki fonemik varyasyonları araştırmak için bir çalışma tasarlanmıştır. Varyasyonları bulmak için Urduca ve Türkçede yaygın olan 75 kelimelik bir liste metin halinde hazırlandı. Bu çalışmanın verileri Lahor'daki Pak-Türk okulundan toplanmıştır. Hazırlanan kelime listesi 10 Türkçe konuşan kişiye verildi. Bu konuşmacılar amaçlı örnekleme tekniği kullanılarak seçilmiştir. Konuşmacıların sesleri kaydedildi ve standart Türk IPA sembollerine dönüştürüldü. Levenshtein algoritma çerçevesi, Türkçe fonemik transkriptlerinin aynı kelimelerin standart Urduca fonemik transkripsiyonları ile karşılaştırmalı bir analizini yapmak için kullanılır. Levenshtein algoritması yardımıyla her iki dilde benzer kelimelerdeki fonemik farklılıklar ölçüldü. Çalışmanın ilgi çekici sonucu, her iki dili konuşanların kullandığı benzer kelimelerin üretim yönlerini anlamaya yardımcı olacaktır.

Anahtar sözcükler: Urdu dili; Türk Dili; sesbirim varyasyonları; Levenshtein algoritması

AUTHOR BIODATA

Tania Ali Khan works as a lecturer at Mihaj University Lahore, Pakistan. Her research interests are morphology, syntax, phonetics and phonology, and language teaching.