Identification of Diphthongs in Urdu and their Acoustic Properties

Rashida Bhatti

Benazir Mumtaz

Sarmad Hussain

Center for Language Engineering,
Al-Khawarizmi Institute of Computer Science,
University of Engineering and Technology, Lahore
firstname.lastname@kics.edu.pk

Agenda

- Introduction
- Need of present research
- Review of the literature
- Methodology
- Results
- Recommendations

Introduction

Dichotomy of Diphthong Definition

- "Diphthongs are single vowels with continuously changing qualities".

Ladefoged (1982: 171)

- "A diphthong may be defined as a combination of two perceptually different vowel sounds within one or same syllable".

Catford (1977:215)

Cont…

- Diphthong is a single sound produced when two vowels (one dominant in duration and stress, one reduced in duration and stress) are paired together in a sequence. House (1998)

- A diphthong being a combination of two vowels has three critical points where their formants present some meaningful information. These critical points are:
  - On glide
  - Transition phase
  - Off glide

Kent and Read (1992)
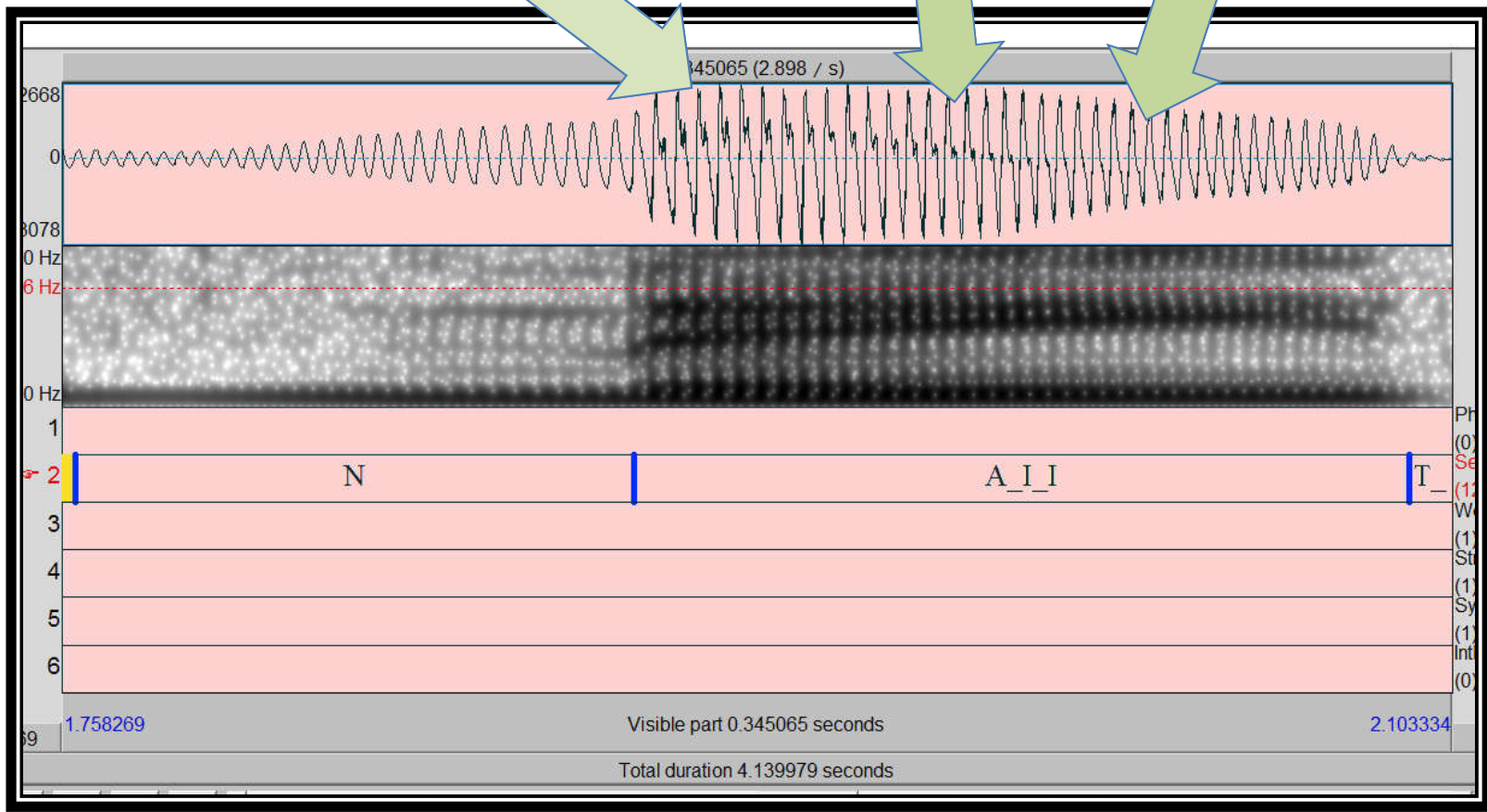
Cont...

- Urdu has a large count of vowels in its Phonetic Inventory.
  - 7 Long vowels, 7 Long nasalized vowels, 3 Short vowels, 3 short nasalized vowels,3 Medial (Majhul)
- It creates the possibility of diphthongs.
- Existence of the phonetic diphthongs is accepted but it is not documented yet.

Review of the Literature

- No evidence of phonemic diphthongs in standard Urdu language but in some Indian local dialects and accents of Urdu has phonemic diphthongs.

Khan(1997)

 – Phonemic diphthongs
  - The meaning will be changed in a particular word if only the vowel onglide is produced. Therefore, the vowel is realized as a monophthong. e.g Mouse/mɑʊs/ and moss /mɑs/.
 – Phonetic/non phonemic diphthongs
  - The meaning would not change in a particular word if the vowel were to be pronounced as a monophthong [ɑ:jɑ:] versus a diphthong [ɑ:æɑ:].

Diphthongs in other Languages

- English:
  - 8 diphthongs
  - i.e. /aɪ/,/eɪ/, /ɔɪ/, /ɪə/,/eə/, /ʊə/, /əʊ/,and /aʊ/

    Roach (1998)

- Persian:
  - No diphthong in Iranian Persian

    Ganjavi et al. (2003)

  - Two diphthongs /ej/ and /ow/
  - Not mentioned either phonemic or phonetic

    Yaesoubi (2010)

Cont…

- Punjabi :
  - 8 diphthongs
  - i.e. /ɪə/, /ɪo/,/ɪə/, /ɪɑ/, /ʊɑ /,/əɪ/, /əe/and /əu/

    Bhatia (1993)

- Hindi :
  - Vowel sequences [əi] and [əu] as:
    - Monophthongs in standard Hindi
    - Diphthongs in Eastern dialects and many Western dialects of Hindi

    Shapiro (2003)

Cont…..

- Urdu:
  - To date four studies
    - 13 diphthongs by Waqar & Waqar (2002) i.e. /ɪũ:/, /əe:/, /əi:/, /ɑ:ɔ:/, /ɑ:i:/, /ɑ:e:/, /e:a:/, /o:i:/, /ɑ:ẽ:/, /o:e:/, /əĩ:/, /ɪa:/ and /a:ũ:/
    - 18 diphthongs by Khurshid, Usman and Butt (2003) i.e. /oi/, /oe/, /ɪo/, /əi:/, /əe:/, /ua:/, /uə/,/a:ɪ/,/ao/, aũ:/, /ɪũ:/, /io/, /ea/, /eo/, /ʊa/, /ui/ and /ue/

Cont…..

- – 17 diphthongs by Sarwar, Ahmad and Tarar (2003) i.e. /ai/, /ae/, /ao/, /ɪũ/, /ɪa/, /aũ/, /oi/, /oe/, /oĩ/, /əi/, /aẽ/, /ea/, /əĩ/, /ʊa/, /ui/, /ue/ and /əe/

- – 7 diphthongs by CLE researchers i.e. /əi/, /əe/, /ae/, /ai/, /ɪũ/, /æa/ and /ʊi/

Need of Present Research

- Unified list of diphthongs
- Addition in the phonetic inventory of Urdu language
- Smooth annotation of Urdu speech corpus for the development of human assisting tools(e.g. ASR, TTS, Screen reader, etc).
- Resolve the issues faced during annotation of speech corpus on multi tiers

Research Methodology

- Combined list of 26 diphthongs(24 given in previous studies and 2 in present research)

- Corpus: Carrier sentences containing 78 words (3*26=78) from 6 speakers(3 males 3 females)

- Two approaches are used in this study i.e. perceptual approach and acoustic approach respectively.

Cont...

- Perceptual approach: Diphthongs are identified by **ten** native speakers using syllable identification technique.

  – Diphthongs which passed the perceptual test were sent forward for acoustic testing.

- Acoustic approach, speech of **six** native speakers is analyzed using durational (both in stressed and unstressed form)and formant cues (F1, F2 and F3) on PRAAT.

Corpus development and speech analysis

- Sample and size of speech corpus :
  - Recordings from 6 native speakers of Urdu (3 males & 3 females), in mono form, at the sampling rate of 48 KHZ [3 instances for each sentence]
  - Ten sentences containing distracters were also recoded
- Software for recording and annotation:
  - PRAAT
- Phonetic character set:
  - Case Insensitive Speech Assessment Method Phonetic Alphabets (CISAMPA)

Perceptual Experiment Methodology

- Three utterances of each diphthong were selected (26*3*6=468), distracter sentences were also added.

- Ten native speakers (5 males and 5 females) were asked to listen and count syllables in words.

- 16 diphthongs were finalised in this experiment.

Sr No.	Diphthongs	Perceptual Agreed Diphthongs	Perceptual Disagreed Diphthongs
1	ɑ:e:	90 %	10 %
2	ɑ:e	80 %	20 %
3	ɑ:ẽ:	70 %	30 %
4	ɑ:ɪ	90 %	10 %
5	ɑ:i:	80 %	20 %
6	ɑ:ĩ:	60 %	40 %
7	ɑ:o:	80 %	20 %
8	ɑ:u:	40 %	60 %
9	əe	100 %	0 %
10	æɑ:	90 %	10 %

11	əi	100 %	0 %
12	əĩ:	80 %	20 %
13	eɑ:	70 %	30 %
14	e:o:	20 %	80 %
15	ɪɑ:	10 %	90 %
16	ɪã:	30 %	70 %
17	ɪo:	20 %	80 %
18	ɪõ:	80 %	20 %
19	ɪu:	0 %	100 %
20	ɪũ:	90 %	10 %

21	o:e:	50 %	50 %
22	o:i:	80 %	20 %
23	o:ĩ:	50 %	50 %
24	uɑ:	30 %	70 %
25	ue:	20 %	80 %
26	ui:	80 %	20 %

Red colored are proposed diphthongs in present research

Acoustic Experiment Methodology

- To verify the proposed list of 16 diphthongs, durations of finalized diphthongs and formant frequencies are analyzed manually.

- Duration of diphthongs in both stressed and unstressed forms are calculated separately (3 unstressed+3 stressed*16 diphthongs*6 speakers=576).

- Only one perceptually selected diphthong /ɑːĩː/ is rejected at this stage of experimentation.

Cont...

- To measure the formant frequency of first (F1), second (F2) and third formant (F3), diphthongs were divided into three components, i.e. on glide (1), transition (2) and off glide (3).

  - F1, F2 and F3 are measured manually from the middle of component 1 and 3.

  - Window of PRAAT was assured to be 20 ms to take formant values of each component.

  - Three instances of every diphthong from the recorded speech of six speakers (3*15*6=270) are considered for formant values.

RESULTS

Perceptual Analysis

- Selection of diphthongs was done on the basis of frequency of the responses.

- In this research, a diphthong is considered a diphthong only, if 70% votes are in favor of a particular diphthong.

- Only one diphthong /ɑːĩː/ having 60% votes was selected for further testing.

Duration Analysis

- Duration analysis shows that diphthongs show the qualities as an entity like long vowels.
- The duration of diphthongs increase in the state of stressed syllable like long vowels. On average the maximum durations of unstressed diphthongs is. 148 ms.
- Therefore, on the basis of durations, /ɑːĩː/ diphthong was rejected at acoustic experiment stage (300-399ms).
- Obtained results show that there is no significance difference in average duration values of diphthongs on the basis of gender.

Cont...

- During annotation and perceptual analysis of diphthongs, it is also observed that in diphthong both vowels blend in such a way that listeners cannot separate them (Figure 2).
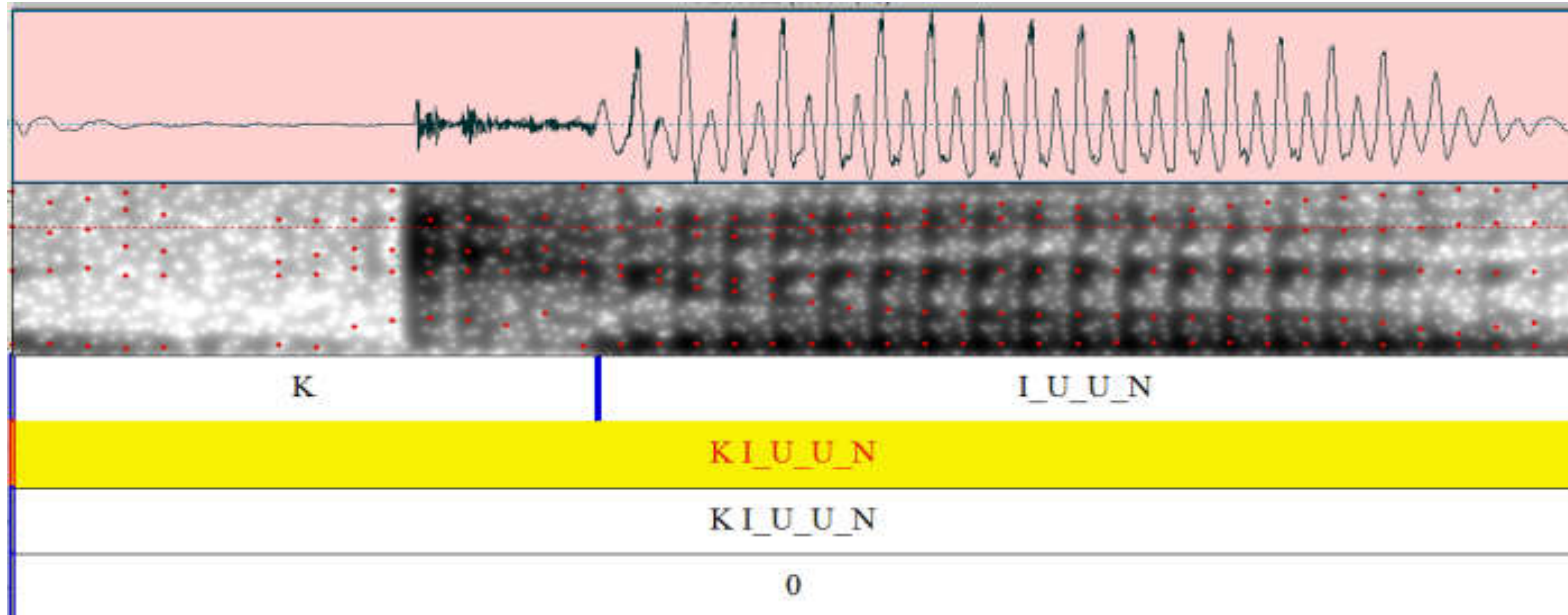
Formant Analysis

- Formant analysis of finalized 15 diphthongs shows that F1, F2 and F3 of vowel components in different diphthong combinations do not have much difference in values.

- One vowel shows almost similar values in different combinations of diphthongs. For example /ɑ:/ has not much difference in different combinations of diphthongs like in /ɑ:e:/, /ɑ:e/, /ɑ:ẽ:/, /ɑ:i:/, /ɑ:ɪ/ and /ɑ:o:/

Cont...

- Obtained results show that although vowel maintains their qualities and formant frequency but they blend with other component to become a diphthong.

- Moreover, analysis of Urdu diphthong highlights that similar to other languages, Urdu diphthongs have three components, i.e. first vowel, transition period and second vowel.

Cont…

- Data analysis of selected diphthongs suggests that diphthongs in Urdu can occur in 5 types of combinations.
    - Short and long oral vowels
        - /ɑːɪ/, /əeː/ and /əiː/
    - Long and long oral vowels
        - /ɑːeː/, /ɑːiː/, /ɑːoː/, /oːiː/and /uːiː/
    - Medial and long oral vowels
        - /ɑːe/, /æɑː/ and /eɑː/
    - Short and long nasalized vowels
        - /əĩː/,/ɪõː/ and /ɪũː/
    - Long and long nasalized vowels
        - /ɑːẽː/

Cont…..

- Combination of long-long vowel and long-medial vowel is unique property of Urdu language.

- Urdu speakers alternate the schwa and J with medial vowel /æ/. The formants of medial vowel /æ/ in /æɑ:/ diphthong and individually has almost similar values (Table 1).

Table 1

Average Formant Value	F1 Hz	F2 Hz	F3 Hz
æ in Diphthong /æɑ:/	575	214	299
Cardinal vowel æ	567	226	309

Cont…..

- In the case of diphthong /eɑ:/ (e.g. in زیاده /zeɑ:ḍɑ:/ excessive, تیاری /ṯeɑ:ri:/ preparation and فلکیات/falkeɑ:ṯ/ universe), Urdu speakers alternate the sound /ɪ/ and /j/ with medial /e/. This medial vowel blends with the following vowel /ɑ:/ and makes the diphthong /eɑ:/ as shown in Figure 1.

- All the speakers pronounced it as diphthong and 70% respondents recognized it as diphthong.

Figure 1
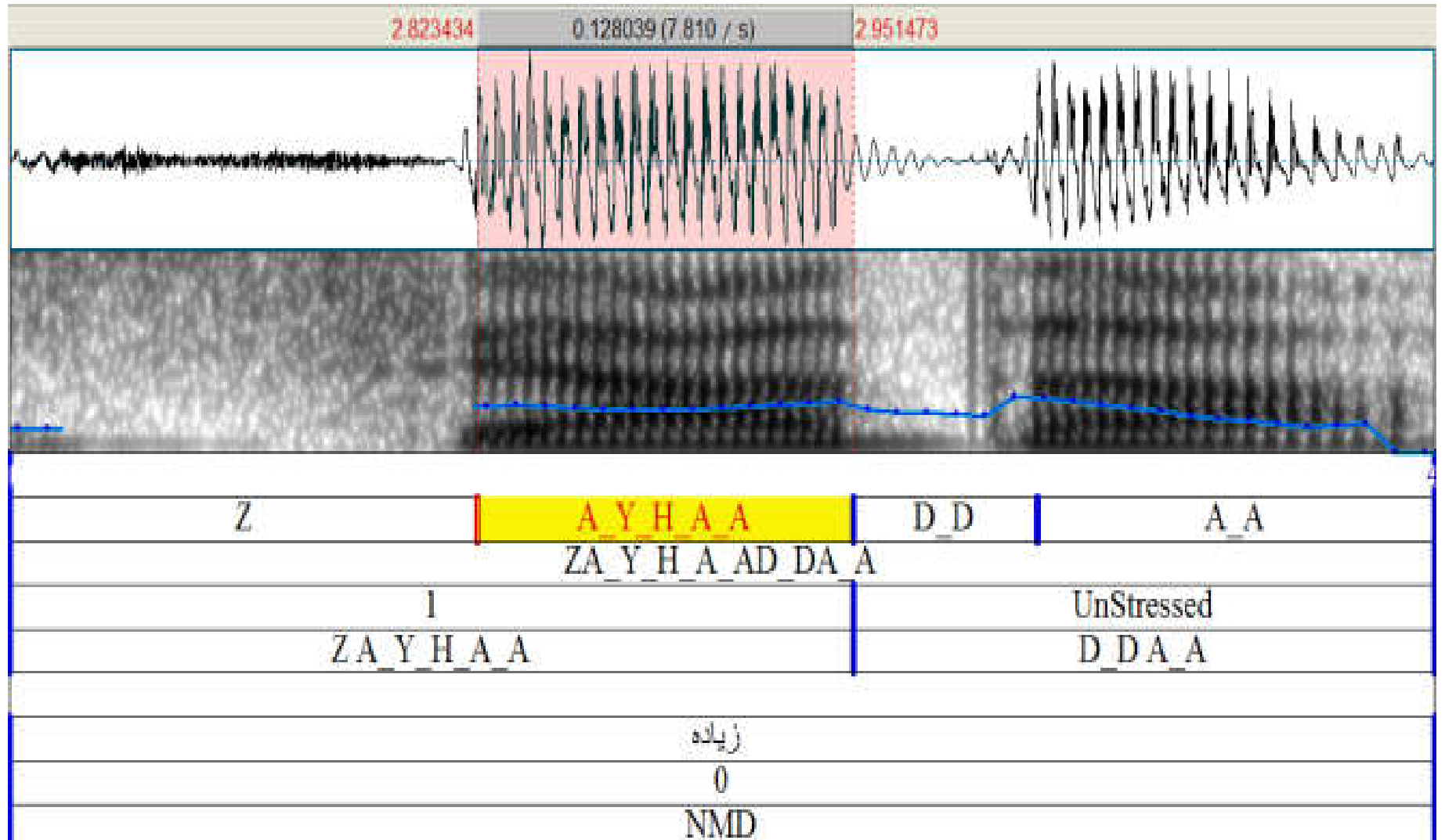
Future work and recommendations

- The selected diphthongs can be added in Urdu phonetic inventory.
- It will be helpful in maintaining the accuracy and consistency during the annotation of speech corpus.
- By marking diphthongs, syllables and stress tier annotation can also be done more smoothly and accurately.
- Moreover, pronunciation lexicon can become more robust using list of diphthongs.
- Hence, Urdu speech database would be more accurate and will represent the quality speech of native speakers.

Cont…

- Moreover, this study reports that the sounds /ə/ and /j/ replace with /æ/ and the sounds /ɪ/ and /j/ replace with medial /e/ to form diphthongs.

- These alternation results are based on the speech of six native speakers. This phenomenon needs to be studied on a large sample to confirm the trend of vowel shifting or alternation among the native Urdu speakers.

References

- B. Mumtaz, A. Hussain, S. Hussain, A. Mehmood., R. Bhatti, M. Farooq, & S. Rauf. "Multitier Annotation of Urdu Speech Corpus", Conference on Language and Technology (CLT14), Karachi, Pakistan, 2014.
- L. Swaran, "Challenges for Design of Pronunciation Lexicon Specification (PLS) for Punjabi Language", Department of Information Technology, Govt of India, 2011.
- J. C. Catford, "Fundamental Problems in Phonetics", Edinburgh: Edinburgh University Press, 1977, pp 215.
- P. Ladefoged,. "A Course in Phonetics", Los Angeles: University of California, ed 2, 1982, pp 171.
- M. A. Khan, "Urdu Ka Sauti Nizam", National Language Authority, Pakistan, 1997.
- A. Q. Khan, "A Preliminary Study of Pahari and Its Sound System", The criterion An International Journal in English Vol IV, Issue IV, 2013, pp 1-20.
- A. Waqar, & S. Waqar, "Identification of Diphthongs in Urdu and Their Acoustic Properties", 2002.
- F. Alam, S. M. Habib & M. Khan, "Acoustic analysis of Bangla vowel inventory" BRAC University, 2008.
- R. D. Kent, C. Read, & R. D. Kent, "The acoustic analysis of speech", (Vol. 58). San Diego: Singular Publishing Group, 1992.
- F. Okati, P. Helgason, & C. Jahani, "Diphthongization in Five Iranian Balochi Dialects", Orientalia Suecana, Vol 61, 2013, pp 107-119.
- C. P. Prinsloo, "A comparative acoustic analysis of the long vowels and diphthongs of Afrikaans and South African English", 2006.

Cont…

- B. Comrie, "Languages of the world. *The Handbook of Linguistics*", 2001, pp 522 (online book: cited on May 09,2016).
- C. P. Masica, "*The Indo-Aryan Languages*", Cambridge University Press, 1993.
- Samare, Yadollah (Samare, Yadollāh). "*Āvāšenāsi-ye zabān-e fārsi, āvāhā va sāxt-eāvāyi-ye heǰā*", Tehran: Markaz-e Našr-e Dānešgāhi, 1368 [1989–90].
- S. Ganjavi, P. G. Georgiou, & S. Narayanan, "ASCII based transcription systems for languages with the Arabic script: The case of Persian", 2003, In Automatic Speech Recognition and Understanding, 2003. ASRU'03. 2003 IEEE Workshop on IEEE, pp 595-600.
- M. Yaesoubi, "*Machine Transcription Conversion between Perso-Arabic and Romanized Writing Systems*", Master thesis, Linköpings University, 2010. Online at: http://liu.divaportal.org/smash/record.jsf?pid=divas2:360341 [Retrieved May 30, 2016].
- A. Hakimi, "Comparative Phonetic Study of Frequently Used Words in Iranian Farsi versus Tajik Farsi", *Journal of American Science* 8(4), 2012, pp 6–16.
- K. Khurshid, S.A. Usman, & N. J. Butt, "Possibility of existence and identification of diphthongsand triphthongs in Urdu language". Akhbar-e-Urdu, 2003, pp 16-21.
- A. Sarwar, S. Ahmed, & A. A. Tarar, "Diphthongs in Urdu language and analysis of their acoustic properties", Annual Student Report, Center for Research in Urdu Language Processing (CRULP), 2003, pp 9-15.

Analysis

- Analysis A (Duration Analysis)
- Analysis B (Formant Analysis)

Analysis A

Sr No.	Diphthong	Average Duration Males (ms)	Average Duration Females (ms)	Minimum Duration males in Unstressed (ms)	Minimum Duration females in Unstressed(ms)	maximum Duration males in stressed (ms)	maximum Duration females in stressed (ms)
1	ɑ:e:	254	207	127	147	296	267
2	ɑ:e	231	212	186	169	280	280
3	ɑ:ẽ:	282	212	200	170	277	276
4	ɑ: ɪ	212	175	182	134	288	228
5	ɑ:i:	233	205	205	150	261	262
6	ɑ:o:	205	221	127	179	268	286
7	əe:	236	186	177	153	272	226
8	æɑ:	321	215	162	148	313	342
9	əi:	234	186	176	153	302	226
10	əĩ:	235	206	201	147	289	258

11	eɑ:	207	195	142	142	241	271
12	ɪõ:	203	152	179	146	225	165
13	ɪũ:	227	126	160	147	273	209
14	o:i:	244	192	229	96	267	293
15	ui:	201	210	167	161	231	327
16	ɑ:ĩ:	300	285	350	336	322	354

Analysis B

Sr No.	Diphthong	First Component			Second Component		
		Average F1 Hz	Average F2 Hz	Average F3 Hz	Average F1 Hz	Average F2 Hz	Average F3 Hz
1	ɑ:e:	813	1707	2990	501	2407	2977
2	ɑ:e	797	1376	2700	542	2035	2970
3	ɑ:ẽ:	890	1534	3144	716	2221	3126
4	ɑ: ɪ	807	1527	2832	531	2154	2928
5	ɑ:i:	788	1694	2792	330	2379	3017
6	ɑ:o:	712	1431	2884	612	1184	2956
7	əe:	454	1830	2631	422	2187	2752

8	æɑ:	525	2048	2759	746	1541	2622
9	əi:	517	1835	2682	319	2477	3094
10	əĩ:	418	2228	3039	332	2118	3046
11	eɑ:	419	2103	2763	683	1492	2582
12	ɪõ:	351	2104	2893	553	1412	2827
13	ɪũ:	331	1637	2596	347	1146	2732
14	o:i:	431	1212	2782	332	2360	2889
15	ui:	313	1735	2708	295	2464	2974

Thank You