# ACOUSTIC ANALYSIS OF DIPHTHONGS IN SINDHI

Ayaz Keerio Muhammad Zeeshan Patoli Bhargav Kumar Mitra Chris Chatwin Rupert Young Philip Birch

Industrial Informatics and Manufacturing Systems Research Centre, School of Engineering and Design, University of Sussex Falmer, Brighton BN1 9QT

#### ABSTRACT

Sindhi is one of the official languages spoken in the Sindh province of Pakistan, and in some parts of India. This paper aims to determine the existence of diphthongs in the language inventory. Acoustic analysis has been carried out for twenty seven diphthong words selected in two different ways: (i) (diphthong-) words in isolated utterances and (ii) (diphthong-) words embedded in a sentence (carrier sentence), read by a speaker. The voice-samples were collected from ten native male speakers from the Sindh region in Pakistan. The main acoustic parameters tested for each occurrence are the formant frequencies, fundamental frequency, and the duration.

Keywords Diphthong, Vowel Glide, Formant Frequencies, Falling Diphthongs. Rising Diphthongs

## INTRODUCTION

The formation of two consecutive vowel sounds in the oral cavity for one syllable results in a glide from the start-vowel to the end-vowel, usually referred to as a diphthong. Diphthongs are also defined as an independent vowel-glide; which means that the articulators start in the position of one vowel and move in the direction of another vowel (Jones, 1969). Diphthongs are well correlated with the moving pattern of the formant frequency plots, starting at the on-glide section of the plot and ending at the offglide section. For example, this feature can be observed in the English diphthong [a1] which involves plot changes from the onglide [a] segment to the off-glide [1] segment (Kent, 2002). Diphthongs are also considered as vowel sounds which inherit the qualities of two vowels. This can be observed in the diphthong sound [e1], of English, ending with the articulators set for the vowel [1] in the oral cavity that entails tongue movement forward and upward from the position set while pronouncing [e] (Raphael, 2006). The same can be observed for [a1], and [o1], etc. diphthongs of English (Raphael, 2006). On the other hand, the diphthongs [au] and [ou] of English ending with the articulators set for the vowel [u] in the oral cavity entail tongue movement backward and upward simultaneously with lip protrusion (Raphael, 2006). This paper focuses on the acoustic phonetic characteristics of diphthongs in the Sindhi language inventory and also on the relationship between the three types of sounds (diphthongs, glides and oral vowels) of the language. The glides like vowels are the sounds having well defined and predictable formant structure; hence there could be confusion with the identification of these three classes of sounds (glides, diphthongs and oral vowels) (Olive, 1993). There are two glide consonants present in the phonemic inventory of Sindhi /w/, and /y/ (Jennifer, 2006). Glides and diphthongs are referred to as the sounds with gradually changing formant pattern (Ladefoged, 1993). The difference between oral vowels and glides is quite straight forward; the glides do not maintain steady state; whereas the vowels maintain steady state (Kent, 2002), (Raphael, 2006). Unlike the oral vowels the difference between diphthongs and glides can be observed by analysing the transition segment of the two sound types. transition segment of diphthongs is slower and more gradual compared to the vowel-to-glide transition (Olive, 1993). The formation of vocal tract for the production of a glide move in and out very quickly; therefore the glides are considered as sharp transitory vowels (Raphael, 2006). A glide in a syllable if precedes the vowel it forms an independent phoneme of Sindhi; if it follows the vowel and produced with a preceding vowel; it produces the diphthong of Sindhi and phonetically diphthongs are referred to as individual phonemes (Olive, 1993).

One of the early study talks about the diphthongs of Sindhi is Ernest Trump's book, 'Grammar of the Sindhi Language', in which author has mentioned that there are no diphthongs in Sindhi; the vowels [a1], and [a0] are loosely pronounced as a-1 and a-0

(Trumpp 1970).

The subject of diphthongs in Sindhi became contentious in the literature, when Jatoi, in the book 'Ilm Lisan and Sindhi Zuban', (written in Sindhi script), reported that diphthongs are present in Sindhi. For example the diphthong [a1], in word 'l[a1]', and the diphthong [iu], in word 'p[iu]' etc (Jatoi, 1996). Bhugio discussed the presence of two diphthongs in Sindhi: (i) [a1] a nonlabialised descending diphthong with a glide from [a] to [1] and (ii) [au] a labialised descending diphthong with a glide from [a] to [v] (Bhugio, 2006). The presence of two diphthongs discussed by (Bhugio, 2006) is also discussed by (Jennifer, 2006) author has associated the diphthong realization in Sindhi with the speaker's religious affiliation such as Muslim Sindhi speakers and Hindu Sindhi speakers. The Muslim Sindhi speakers produce the short diphthongs [a1], and [a0], whereas the lax vowels [8] and [9] are typical for Hindu Sindhi speakers. The glides of Sindhi /w/ and /y/ (Jennifer, 2006), at word initial position, function independent phoneme and at word final or medial position, if followed or preceded by a consonant, they function as vocalic glides. The glides in Sindhi, if preceded or followed by a vowel, at word final or medial position the two sounds produced together are considered as a diphthong phoneme of Sindhi (http://tdil.mit.gov.in /sindhi-designguideoct02.pdf). The two view points can be concluded from the literature regarding the presence of diphthongs in Sindhi. First when a short vowel with 'hamzo', in a syllable or at the word final or medial position, if followed or preceded by a vowel, could potentially form a diphthong of Sindhi. Second, if a glide is preceded or followed by a vowel at the word final or medial position can potentially form a diphthong of Sindhi. In order to carryout acoustic phonetic analysis of Sindhi diphthongs according to the above mentioned two perspectives, the selection of the words were made in a way so that these words contain the syllable(s) having a short vowel with 'hamzo', either preceded or

followed by a vowel or the syllables containing the glide followed or preceded by a vowel. A syllable in Sindhi must have a vowel that serves as a syllable nucleus; the consonants seldom in Sindhi can not form a syllable. The two adjacent vowels in the syllable of Sindhi are paired as long and short vowels. The syllables of the following structure can be constructed in Sindhi (Jatoi, 1996):

- V: only long vowel individually can construct a syllable
- Cv: consonant followed a short vowel
- CV: consonant followed by a long vowel
- CVv: consonant followed by two successive vowels; in this syllable the final vowel is a short vowel
- CVC: a vowel surrounded by consonants
- CCV: two consecutive consonants followed by a vowel
- CCVC: two consecutive consonants followed by a vowel and a consonant
- CVCC: syllable ends with two consecutive consonants Note: symbol C = consonant, V = long vowel, and v = short vowel.

The need to carry out more research work on the subject of diphthongs in Sindhi is highly desired; even though it is to be assumed that diphthongs do exist in Sindhi, their identification is vet to be registered and type categorized. This paper aims to identify and measure the acoustic properties of Sindhi diphthongs. In order to measure the acoustic properties of diphthongs in Sindhi, we will be focusing the word utterances containing the syllable(s) of the CV, CV and CVv structure and the mono-syllabic words formed by these syllables for the purpose of the voice samples recording.

The rest of paper is organized as follows: section II describes the methodology; section III discusses the experiments; section IV presents the acoustic analysis; and finally the conclusions are drawn in section V

## METHODOLOGY

# A. Data collection

Speech samples were recorded from 10 native Sindhi male speakers. The speaker's selection criterion was set as graduate aged 20 to 40 years. A list of twenty seven Sindhi words was made; to the best of our perception that the phoneme sequence of these words may carry a diphthong sound. An equal number of other non-diphthong words were added to the selected word list and randomized. Speakers were given the randomized list of the isolated words and the list of words embedded in a carrier sentence. Speakers were asked to read out the two word lists for the voice sample recording. Ten instances of each speaker's voice samples were recorded and stored as .wav file format for subsequent offline processing. Only one speaker voice samples were recorded at one time to avoid any pronunciation imitation. The speech samples were recorded in the laboratory keeping the laboratory climatizer on but doors and windows closed.

B. Tools used for data collection

For recording a portable solid state digital audio recorder (Marantz PMD660) was used with an external broadcast quality lightweight condenser microphone (the RODE NTG-2). The Speech samples were recorded as a single channel with a sampling rate of 48 kHz, and stored in wav file format, 16 bits per sample. On an average the mouth to microphone distance was maintained at four inches.

## EXPERIMENTS

Two experiments were set for the acoustic analysis of the language diphthongs: (i) analysis based on the speech samples of the isolated word utterances and (ii) the analysis based on the speech samples recorded as a carrier sentence containing the target diphthong word from the list of twenty seven selected words. In this way the statistical analysis results of the entire study are based on the 5400 [= 2 x (10 x 27 x 10)], voice samples captured from the native Sindhi speakers. The reason to carry out analysis in two ways is that there exist the temporal differences, when speaker speaks out isolated words compared to the words embedded in a carrier sentence. The isolated word utterances in this work are considered as an ideal case for the acoustic analysis; because the vocal tract configurations are ideally set for the target word utterance.

#### ACOUSTIC ANALYSIS

A diphthong is a combination of two vocalic elements with three significant analysis points: (i) on-glide (the duration of the first vocalic element) (ii) transition (the duration of the shift between first and second vocalic elements), and (iii) off-glide (the duration of the second vocalic element), (Gay, 1968). In order to quantify the acoustic characteristics of diphthongs in Sindhi; the main parameters measured are the first two formants (F1, F2), fundamental frequency (F0), and the duration. The speech samples of the two experiments were gathered and the samples of the diphthong segment were extracted from the word utterances and analyzed. The diphthong boundaries in the word utterances were marked manually by inspecting the spectrograms of the word utterances. Spectrograms of the corresponding speech samples were computed using the Praat (Boersma, 2009) a speech analysis tool. Table I. below shows the list of the words selected for the acoustic analysis of Sindhi diphthongs.

TABLE - I LIST OF WORDS CONSIDERED FOR ANALYSIS THAT MAY CONTAIN DIPHTHONGS

S.No	IPA Symbol	Transcription	Sindhi
1	ε	επολο, εετυ, εδυ	عينڪ، سَنير، عَيبُ
3	19	siə, hiə	, سييءَ، هِيءَ
3	01	po1, go1	پوء، گوء
4	υi	hui, pitui, sui	هٔني ، پيئي، سُئِي
5	1e	thie, pie	لِتَي، پني
6	a1	laı, paı, bud <sup>h</sup> aı	لاء، پاء، بُدّاء
7	uı	khushbu1,	خۇشبۇء
8	Э	fovendo, bho,	چَوندنو، ڀَٺو
9	uə	huə	هُوءَ،
10	αυ	mao, pao	ماء، پاءُ
11	io	sio, pio, dio	سِيءُ، پِيءُ ، جِيءُ
12	19	d <sup>h</sup> 1ə	ڒؠءؘ
13	ບວ	ჭსაი, suar	سُلْر، جُنْر
14	UO	dhuo, puo	ڏنو, پُنو

Note that the diphthongs listed in this study may not be the only diphthongs present in the language inventory; because Sindhi is considered as a flexible language in terms of altering the word pronunciation with a vowel insertion or digestion in between two consonants i.e. the word hikto "one" and the word  $f^hup ti$  "shack" can be alternatively pronounced as hiktoto and  $f^hup ti$  by inserting the short vowels [a], and [1] respectively (Jennifer, 2006). The other factor is the rate of the speaker speech varies from speaker to speaker; in that situation the slow utterance of the glides may result in a diphthong sound (Raphael, 2006).

The first two formants provide fundamental cues for the perception of the monophthong vowels; however the second formant is considered one of the most fundamental acoustic cue for the perception of diphthongs (Manrique, 1979), (Aiza, 2004), (Jha, 1985). The acoustic analysis and the conclusions based on the obtained results were drawn keeping in view the following

characteristics of Sindhi diphthongs in general:

 the diphthongs should have two vocalic targets (Ladefoged, 2001)

• the type of the diphthong falling or rising

- the transition between two vocalic elements is not a hiatus
- the transition between two vocalic elements is a diphthong transition or it is a vowel-to-glide transition.

To accomplish the above mentioned characteristics of diphthongs a set of acoustic properties for each diphthong were measured based on the list of twenty seven words spoken by ten male speakers of Sindhi. The formant frequencies for the first five formants (F1, F2, F3, F4 and F5) were obtained using the Burg LPC (linear predictive coding) algorithm implemented in the Praat (Boersma, 2009) a speech analysis tool. To determine the two vocalic elements present in the diphthong the z-scores based on 5400 samples for the first two formants were computed according to the method described by (Hongyan, 2007), (Lobanov, 1971). The z-scores for first two formants of each diphthong were computed by subtracting the mean of the first formant from the raw F1 values and the mean of the second formant from the raw F2 values. The differences were then divided by the standard

deviation of the raw formants values. The reason for computing the z-scores of the diphthong formants was to demarcate the two vocalic elements of the diphthongs i.e the obtained negative (-ve) z-scores belong to one vocalic element and the positive (+ve) z-scores belong to the other vocalic element of the diphthong. If all the obtained z-scores are either +ve or -ve or they are very irregular, (say +ve, -ve and again +ve, -ve and vice versa), then the phoneme is considered a non-diphthong phoneme of Sindhi. Figures 1 to 7 below show the computed z-scores for the diphthongs of Sindhi.

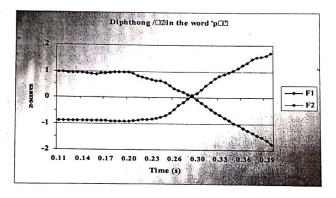


Figure-1 shows the typical z-scores of first two formants of the diphthong [ $\alpha I$ ] present in the word utterance 'p[ $\alpha I$ ]', of Sindhi. Note that F1 starts with +ve z-scores and moves gradually to the –ve z-scores and in similar way the F2 z-scores move gradually from –ve to +ve z-scores

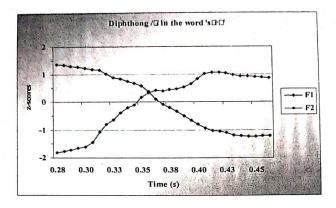


Figure-2 shows the typical z-scores of first two formants of the diphthong [\varepsilon] present in the word utterance's[\varepsilon]ro', of Sindhi. Note that F1 starts with +ve z-scores and moves gradually to the ve z-scores and in similar way the F2 z-scores move gradually from -ve to +ve z-scores

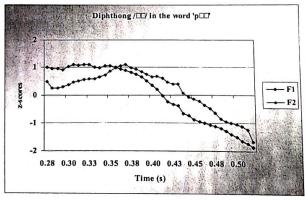


Figure-3 shows the typical z-scores of first two formants of the diphthong [ao] present in the word utterance 'p[ao]', of Sindhi. Note that both F1 and F2 start with +ve z-scores and move gradually to the -ve z-scores

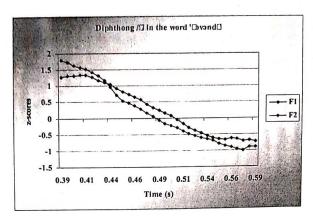


Figure-4 shows the typical z-scores of first two formants of the diphthong [3] present in the word utterance 'favand[3]', of Sindhi. Note that both F1 and F2 start with +ve z-scores and move to the ve z-scores.

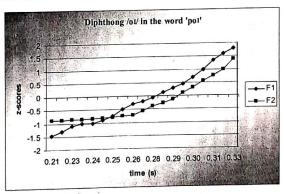


Figure-5 shows the typical z-scores of first two formants of the diphthong [o1] present in the word utterance 'p[o1]', of Sindhi. Note that both F1 and F2 start with -ve z-scores and gradually move to the +ve z-scores

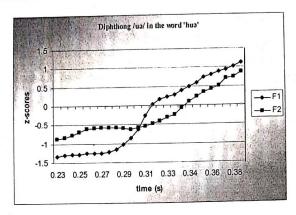


Figure-6 shows the typical z-scores of first two formants of the diphthong [uə] present in the word utterance 'h[uə]', of Sindhi. Note that both F1 and F2 start with -ve z-scores and gradually move to the +ve z-scores

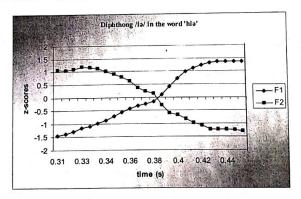


Figure-7 shows the typical z-scores of first two formants of the diphthong [ia] present in the word utterance 'h[ia]', of Sindhi. Note that the F1 starts with -ve z-scores and move gradually to the +ve z-scores and F2 start with +ve z-scores and gradually move to the ve z-scores

Figures 1 and 2 show that the diphthongs with the transition from open back vowel (first vocalic element) to the close front vowel (second vocalic element) yield +ve F1 z-scores for the first vocalic element and -ve z-scores for the second vocalic element and the corresponding F2 z-scores for the first vocalic element are -ve and +ve, respectively, for the second vocalic element. Figures 3 and 4 show that the diphthongs with the transition from open back vowel (first vocalic element) to the close back vowel (second vocalic element) yield +ve F1 and F2 z-scores for the first vocalic element and -ve F1, and F2 z-scores for the second vocalic element Figures 5 and 6 show that the diphthongs with the transition from close back vowel (first vocalic element) to the close front vowel (second vocalic element) yield -ve F1, and F2 z-scores, for the first vocalic element and +ve F1, and F2 z-scores for the second vocalic element. Figure 7 shows that the diphthongs with the transition from close front vowel (first vocalic element) to the open mid vowel (second vocalic element) vield -ve F1, z-scores for the first vocalic element and +ve z-scores for the second vocalic element and the F2 z-scores are +ve for first vocalic element and ve for the second vocalic element.

A significant difference among the z-scores of the diphthongs in Sindhi can be observed. However the sign change from +ve to -ve and vice versa for the first two formants among two vocalic elements in diphthongs is of more importance. If there is a sign transition between the (F1/F2) z-scores of the two vocalic elements (first and second) the presence of a diphthong phoneme in the utterance is guaranteed. Whereas the irregular sign transition of the z-scores among the two vocalic elements confirm the presence of a non-diphthong phoneme in the utterance. In other words more than one F1 and F2 z-score crossovers prove that the phoneme is not a diphthong of Sindhi. Table II below summarizes the obtained analysis results for twenty seven Sindhi words selected for diphthongs analysis.

TABLE - II THE LIST OF WORDS AFTER ACOUSTIC ANALYSIS

Sno	IPA Symbol	Transcription	Sindhi	Diphthong
1	8	snaka, sero, ebo	عينڪ ،سنير ،عيب	Yes
2	iə	siə, hiə	سِيءَ، هِيءَ	Yes
3	01	po1, go1	پوء، گوء	Yes
4	υi	hoi, pitoi, soi	هُني ، پيٽي، سُئِي	No
5	10	thie, pie	ٺِني، پئي	No
6	aı	laı, paı, 60d <sup>h</sup> aı	لاء، پاء، بُداء	Yes
7	uı	khushbu1,	خو شبو ء	Yes
8	Э	favanda, bha,	چَوندنُو، پَئو	Yes
9	uə	huə	الله الله الله الله الله الله الله الله	Yes
10	αυ	maυ, paυ	اءُ، پاءُ	Yes
11	iυ	sio, pio, dio	سِيءُ، پِيءُ ، جِيءُ	Yes
12	19	d <sup>h</sup> ıə	دِيْءَ	No
13	СЭ	фuər,	جُنر ،	No
14	UO	d <sup>h</sup> uo, puo	دُئو	No

Diphthongs are considered moving voicing elements from the beginning vowel to the target vowel without any hiatus (Raphael, 2006), (Gay, 1968). The acoustic F1-F2 vowel plot shown in figure 8 provides the clear picture of the direction in which the formants move gradually in diphthongs. The solid and dashed line arrows in figure 8 below indicate the direction of the formants movement from the first vocalic element (the onset) to the second vocalic element (the offset vowel). The solid line arrows are drawn for the direction of all falling diphthongs from the first vocalic element (the onset) to the second vocalic element (the offset vowel), and the dashed line arrows are drawn for the direction of all rising diphthongs from the first vocalic element (the onset) to the second vocalic element (the offset vowel) of Sindhi. The count of diphthongs in Sindhi shown in figure 8 below may vary; because of the fact that Sindhi is flexible language and the rate of speaker's speech. The slow utterance of the glides and the dialect specific insertion and digestion of the short vowels in Sindhi could potentially lead to the increase or decrease in the count of the language diphthongs. The formant movement of the

diphthongs [a1], [8], and [01] of Sindhi typically starts with the open vowel and ends with the close front vowel as shown in figure 8 and 9. The formant movement of the diphthongs [au] and [b] of Sindhi typically starts with the open vowel and ends with the close back vowel as shown in figure 8 and 10. The formant movement of the diphthongs [iu] and [iə] of Sindhi typically starts with the front close vowel and ends with the low back vowel as shown in figure 8 and 11. It is obvious from the spectrograms of figures 9 to 11, that the formants of Sindhi diphthongs move gradually from one vocalic element (the beginning vowel) to the other vocalic element (the target vowel) without any hiatus or disruption.

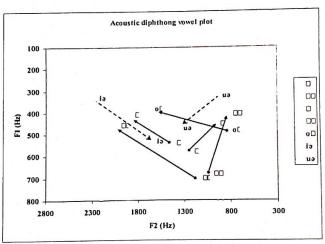


Figure-8 shows F1xF2 vowel plot of Sindhi diphthongs

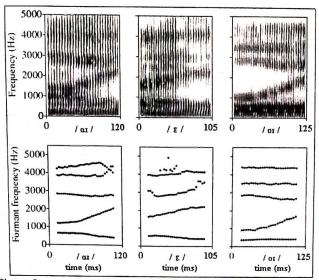


Figure-9 spectrogram of three diphthongs [α1], [ε], and [ο1] of Sindhi along with the formant tracks

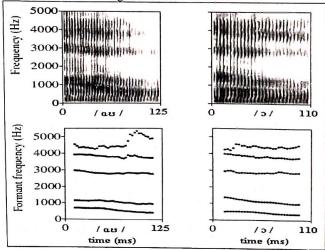


Figure-10 spectrogram of two diphthongs [ao], and [o] of Sindhi along with the formant tracks

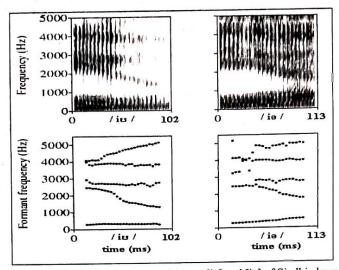


Figure-11 spectrogram of two diphthongs [io] and [iə] of Sindhi along with the formant tracks

#### CONCLUSION

Two experiments were conducted to analyse the diphthong realization in the phonemic inventory of Sindhi. It was observed that the presence of diphthongs in Sindhi mainly depends on the pronunciation and on how fast an individual speaker spoke. The language flexibility of insertion and digestion of the short vowels and the dialect specific pronunciation variation could potentially increase or decrease the count of diphthongs in Sindhi. The obtained acoustic analysis results in this study suggest that when a short vowel immediately follows a long vowel in a syllable or at the word final position the combined production of the two sounds result in a diphthong of Sindhi as in words 'p[iv]', 'h[ia]', 'p[a1]' etc. The only exception is the dialect variation in the word 'l[a1]'; the speakers of Utradi (northern) and Middle pronounce this word with the diphthong [a1] and the speakers of Lari dialect pronounce it alternatively without diphthong as 'l[se]'. The analysis results also show that a short and a long vowel phoneme sequence vV (short vowel followed by a long vowel), in a syllable or in a word utterance do not form a diphthong sound of Sindhi as in words 'p[əi]', and 'khadh[vi]' etc. The diphthongs [ε] and [ɔ] are special in a sense of how speakers are pronouncing them; for example the word's[s]ra', is quite frequently pronounced as 's[e]ra', and the word 's[o]kho', is pronounced as 'sokho', which are phonetically wrong pronunciations of the two sounds and both the words in the latter pronunciation do not form a diphthong sound of Sindhi.

#### REFERENCES

Aiza S., Sana A. and Aymen A.T. (2004), "Diphthongs in Urdu Language and Analysis of their Acoustic Properties", CURLP Annual Student Report

Boersma, Paul & David Weenink (2009), "Praat: doing phonetics by computer (5.1.20)", http://www.praat.org/ (26 October 2009).

Bhugio, M.O. (2006), "The Diachronic Sociolinguistic Situation in Sind", Journal on Cultural Patrimony, vol. 1, January-June.

Gay, T. (1968), "Effects of speaking rate on diphthong formant movements", J. Acoust. Soc. Amer., 44: 1570-1573.

Hongyan, Wang (2007), "Mutual intelligibility of Chinese, Dutch and American speakers of English", Ph.D. dissertation Netherlands Graduate School of Linguistics

Jatoi, Ali. N. (1996), "Sindhi Boli jo Lisani Jaezo", 3rd edn. Sindh Text

Book Board, Jamshoro Pakistan.

Jennifer, Cole (2006), "The Sindhi language", In K. Brown (ed.) Encyclopedia of Language and. Linguistics, 2nd edn, Oxford: Elsevier, v.11: 384-386.

Jha, S. K. (1985), "Acoustic analysis of the Maithili diphthongs" J. Phonetics, 13:107-115

Jones, Danial (1969), "An outline of English phonetics", (reprint). England: W. Heffer & Sons Ltd., Cambridge.

Kent, R. D. & Charles Read (2002), "The acoustic analysis of speech" 2nd edn. Singular Publishing Group.

Ladefoged, Peter (1993), "A Course in Phonetics", 3rd Edition, Harcourt Brace, Los Angeles

Ladefoged, Peter (2001), "Vowels and consonants: an introduction to the sounds of languages", Wiley-Blackwell publishers.

Lobanov B. M. (1971), "Classification of Russian vowels spoken by different speakers", Journal of the Acoustical Society of America, vol. 49(2B), 606-608.

- Manrique A. (1979), "Acoustic analysis of the Spanish diphthongs", Phonetica, 36:194-206.
- Olive. P. J, Greenwood. A., Coleman. J. (1993) "Acoustics of American English Speech: a dynamic approach", Springer-Verlag New York
- Raphael L. J. & Gloria J. B., Katherine S. H. (2006), "Speech science primer: physiology, acoustics, and perception of speech", 5th edn, Lippincott Williams & Wilkins.
- Trumpp P. (1970), "Grammar of the Sindhi Language", edition 1872, reprinted in 1970, Biblio Verlag, Osnabruck