

## COMPARATIVE ANALYSIS OF VOWELS, DIPHTHONGS AND GLIDES OF SINDHI

Ayaz Keerio<sup>1</sup>, Lachhman Das Dhomeja<sup>2</sup>, Asad Ali Shaikh<sup>2</sup>, Yasir Arfat Malkani<sup>1</sup>

<sup>1</sup>Institute of Mathematics and Computer Science, University of Sindh, Jamshoro, Pakistan  
ayazkeerio@hotmail.com, yasir.malkani@usindh.edu.pk

<sup>2</sup>Institute of Information and Communication Technology, University of Sindh,  
Jamshoro, Pakistan  
lachhman123@yahoo.com, asad.shaikh@usindh.edu.pk

### **ABSTRACT**

*Sindhi language is primarily spoken in the Sindh province of Pakistan, and in some parts of India. Languages phonemic inventory include vowels, consonants and diphthongs. This paper presents acoustic analysis and properties of the glide consonants of Sindhi. Glides are considered having stable and predictable formant structure and associated acoustic properties like vowels and diphthongs. Understanding the corresponding acoustic similarities, differences and relationship between three types of these sounds is the subject of discussion of this paper.*

### **KEYWORDS**

*Consonants, Formant frequencies, Glides, Phonemic inventory, Sindhi*

### **1. INTRODUCTION**

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 [1]. There are two glide consonants in the phonemic inventory of Sindhi articulated at two places of articulation: the bilabial /w/, ɤ and the palato-alveolar /y/ ɣ [2]. Glides and diphthongs are referred to as the sounds with gradually changing formant pattern, and the vowels maintain steady-state [3, 4]. The difference between vowels and glides is quite straight forward; the glides do not maintain steady-state; whereas the vowels maintain steady-state [3, 5]. Unlike the vowels the difference between diphthongs and glides can be observed by analysing the transition segment of the two sound classes. The transition segment of diphthongs is slower and more gradual compared to the vowel-to-glide transition [1, 4]. The formation of vocal tract for the production of a glide consonant move in and out very quickly; therefore the glides are considered as sharp transitory vowels [6, 7]. Two phonetic behaviours are associated with the glides /w/ ɤ and /y/ ɣ of Sindhi: (i) at the word initial position or at the word final or medial position, if followed by a consonant, glides function as independent phonemes (ii) the glides if preceded or followed by a vowel, at the word final or medial position and the two sounds are produced together form a diphthong phoneme of Sindhi (<http://tdil.mit.gov.in/sindhidesignguideoct02.pdf>; site visited on 10th October 2010) and phonetically diphthongs are independent phonemes in the phonemic inventory of Sindhi.

Diphthongs are sounds which inherit the qualities of two vowel sounds the start-vowel (on-glide), the end-vowel (off-glide) and the transition segment. The formation of two consecutive vowels 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 [4, 8]. Diphthongs are also well correlated with the moving pattern of the formant frequency plots, starting at the on-glide section of the plot and ending at the off-glide section. For example, this feature can be observed in the English diphthong [aɪ] which involves plot changes from the on-glide [a] segment to the off-glide [ɪ] segment [5]. Glides are vowel-like sounds; but unlike the vowels they do not form the prominent part of the speech signals in a syllable which is why they cannot form syllables individually in Sindhi as vowels do. However this consonant class of sounds possess a few vowel-like acoustic characteristics such as the periodic output waveform and well defined formant structure; due to the well defined formant structure glides are usually referred to as the semi-vowels [9, 10, 11]. Glides are the sounds considered with a gradually changing formant pattern in the same way the pattern changes for the diphthongs [3]. This is why the acoustic analysis of the glide consonants in this paper is carried out together with the analysis of vowels and diphthongs. As discussed above the glides do not maintain a steady-state, this characteristic differentiates them from vowels; however it is difficult to differentiate them from the diphthongs as both sounds gradually change the formant pattern. Glides are sharp transitory semi-vowels; therefore the transition segment duration for these sounds is shorter compared to the transition segment for diphthongs; because the vocal tract configurations during the production of glides move in and out quickly [1, 12]. The difference between glides and vowels is easy to define; because the later class of sounds maintain steady-state. However the difference between glides and diphthongs is difficult to determine; because the two sounds do not maintain steady-state, rather they are the sounds with a gradually changing formant pattern. In this paper we will rely upon the analysis of the transition segment of the glides and diphthongs for their identification and type categorization.

The rest of paper is organized as follows: the section 2 talks about the methodology, section 3 describes glide consonants of Sindhi, section 4 presents acoustic analysis and finally conclusions are drawn in section 5.

## **2. METHODOLOGY**

In order to carryout acoustic-phonetic analysis of the glide consonants of Sindhi the voice samples were collected from 75 native male speakers of Sindhi across five of its regional dialects. Voice sample recordings were made in two ways (i) individual word utterances embedded with target glide consonant in a syllable part of VCV (vowel consonant vowel) phoneme sequence and (ii) words embedded with target glide consonant in a syllable part of VCV (vowel consonant vowel) phoneme sequence in a carrier sentence. Speakers were given the two lists one for the isolated word utterance and the other with the words embedded in a carrier sentence for the recording of the voice samples. Speakers were asked to read out the two lists in their natural style of speaking. At least ten instances of each speaker's voice samples were recorded and stored in .wav file format for subsequent offline processing. Only one speakers voice samples were captured at one time to avoid any pronunciation imitation. Tools used for recording include Marantz (PMD660) a solid state digital audio recorder, 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.

## **3. GLIDES OF SINDHI**

A glide in a syllable if it precedes or follows the vowel it can potentially result in a diphthong phoneme of Sindhi; therefore the voice samples recording for the acoustic analysis of glides

follows the VCV phoneme sequence. In the VCV phoneme sequence the consonant surrounded by vowels is the target analysis glide consonant in the word utterance. This VCV phoneme sequence helps to analyse the formant transition coming into the consonant sound (the glide consonant) from the preceding sonorant sound (the vowel) and the transitions going away from the glide consonant to the following sonorant sound (the vowel). Figure 1 below shows the waveform of the word utterance /həwɑ/ هُوَا containing the bilabial glide /w/ و of Sindhi. The waveform segment for the glide /w/ و shows the periodic signal activity and the presence of the energy in the lower frequency region is shown in the spectrogram of figure 2; this indicates that the bilabial glide /w/ و is a voiced phoneme of Sindhi. The first formant motion due to the low F1 is downward (coming into the glide) and upward (going away from it); whereas the F2 transitions remain unchanged at the vowel junctures for this sound. The glide /w/ و showed sharp transitory formant transitions; therefore the two vocalic elements for the segment of the glide consonant cannot be witnessed in the word utterance /həwɑ/ هُوَا see the spectrogram of figure 2 and formant tracks in figure 3. The acoustic analysis shows that the glide /w/ و of Sindhi, if followed or preceded by a vowel, forms an independent phoneme of Sindhi. Figure 4 shows the waveform of the word utterance /vəyo/ وَيُو containing the palato-alveolar glide /y/ ي of Sindhi. The waveform segment for the glide /y/ ي shows the periodic signal activity and the presence of the energy in the lower frequency region, see the spectrogram of figure 5; this indicates that the palato-alveolar glide /y/ ي is a voiced phoneme of Sindhi. Due to the high F2 the second formant motion is upward (coming into the glide) and downward (going away from it); whereas the F1 motion is downward and upward. The glide /y/ ي of Sindhi shows sharp transitory formant transitions, this is shown in figure 6 the formant tracks. The acoustic analysis shows that the glide /y/ ي of Sindhi, if followed or preceded by a vowel forms an independent phoneme of Sindhi.

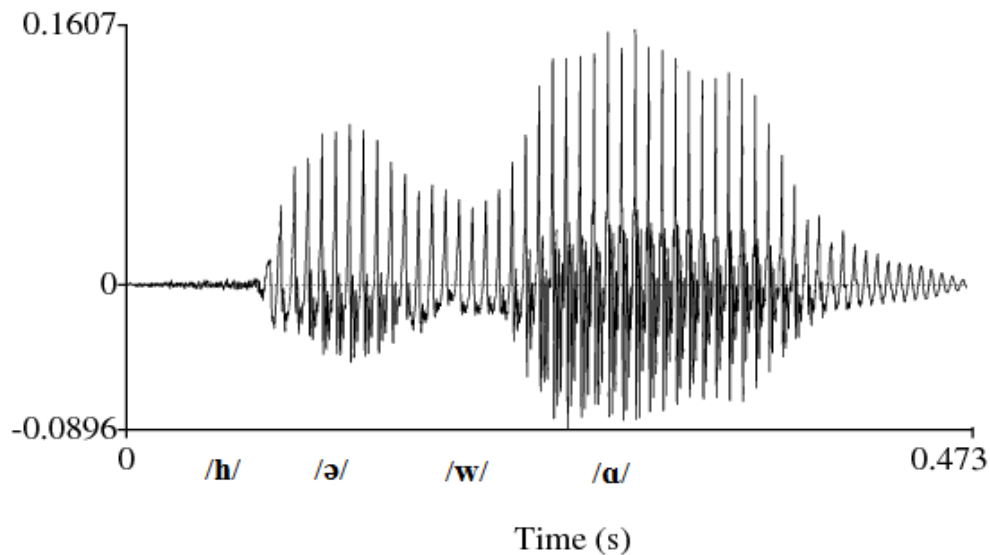


Figure 1. The waveform of the word utterance /həwɑ/ هُوَا containing the bilabial glide /w/ of Sindhi.

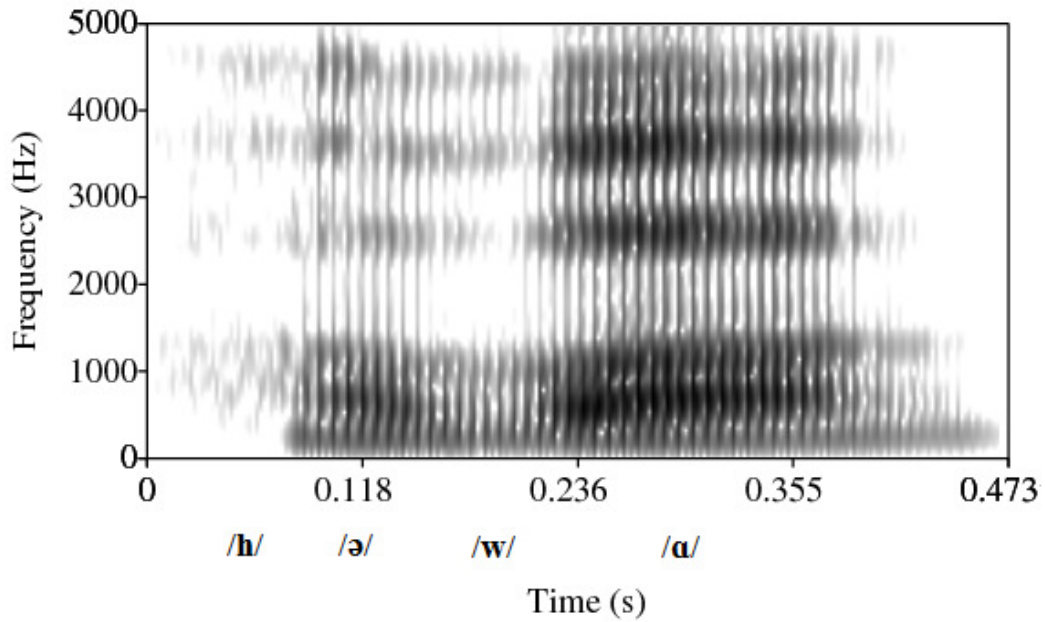


Figure 2. The spectrogram of the word utterance /həwɑ/ هَوَا containing the bilabial glide /w/ of Sindhi.

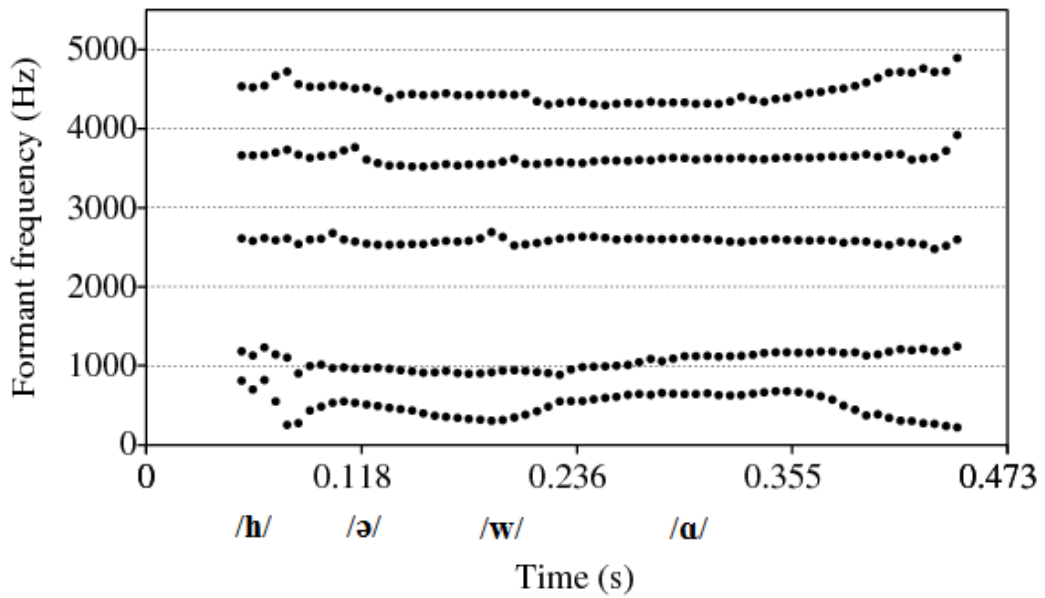


Figure 3. The formant transitions coming into the bilabial glide /w/ from the preceding vowel sound /ə/ and going away from it to the following vowel sound /ɑ/ of Sindhi.

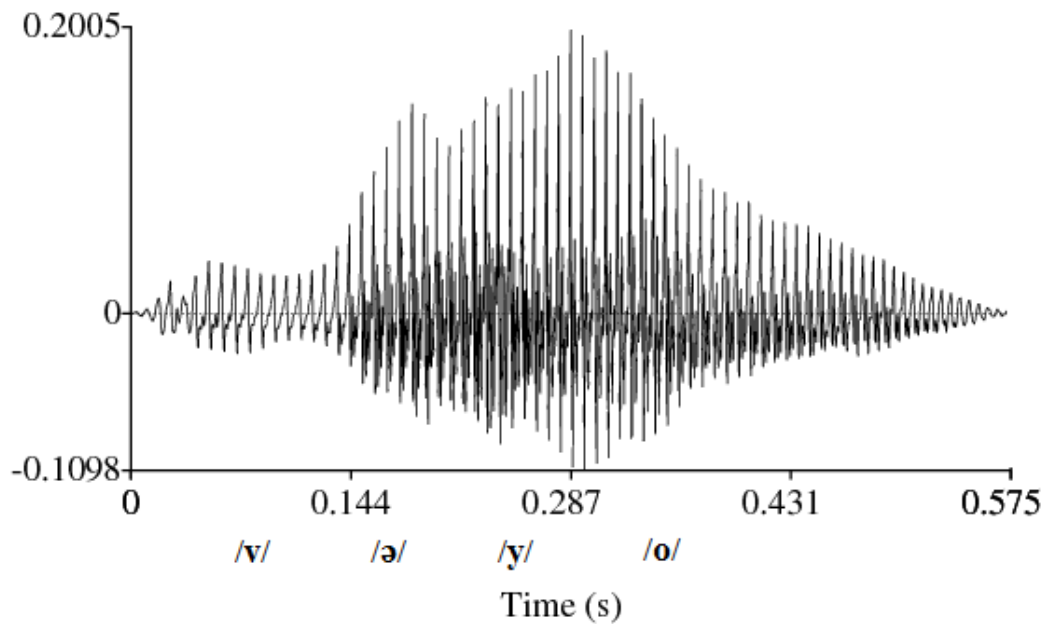


Figure 4. The waveform of the word utterance /vəyo/ ویو , containing the palato-alveolar glide /y/ of Sindhi.

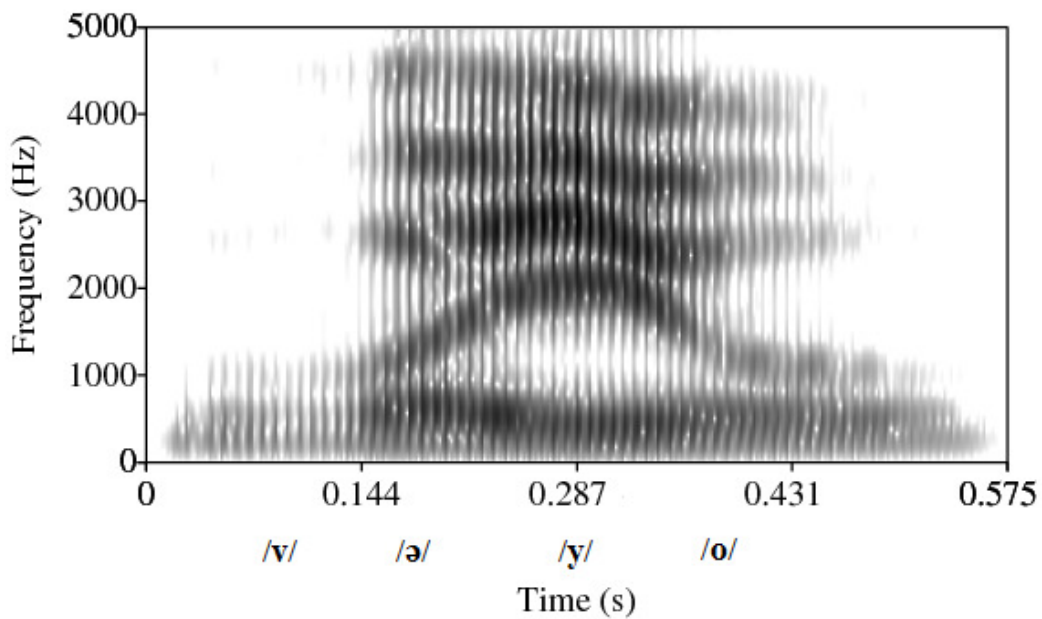


Figure 5. The spectrogram of the word utterance /vəyo/ ویو , containing the palato-alveolar glide /y/ of Sindhi.

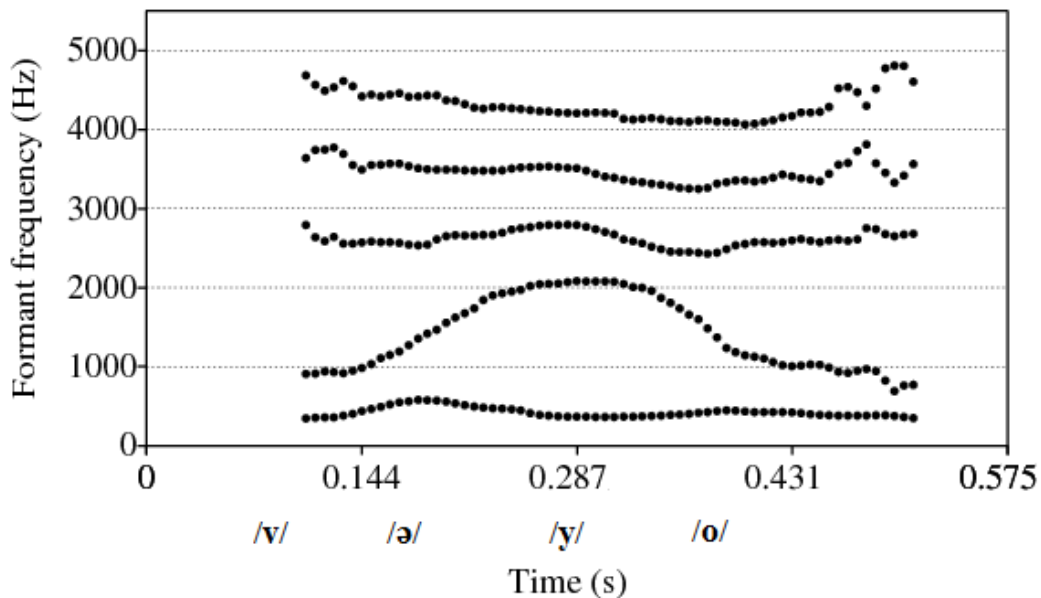


Figure 6. The formant transitions coming into the palato-alveolar glide /y/ from the preceding vowel /ə/ and going away to the following vowel /o/ of Sindhi.

#### 4. ACOUSTIC ANALYSIS

Sounds possess similar acoustic characteristics like glides are the vowels and diphthongs; such similar acoustic characteristics include the periodic output waveforms and the well defined formant structure, with most of energy present in first formant. The only consonant class (the glides) having well defined and predictable formant structure could potentially lead to confusion for the identification of the three classes of sounds: the glides, diphthongs and vowels [1]. Therefore the difference between glides and vowels is easy to define; because the later class of sounds maintain steady-state. However the difference between glides and diphthongs is difficult to determine; because the two sounds do not maintain steady-state, rather they are the sounds with a gradually changing formant pattern. This study relies upon the acoustic analysis of the transition segment of the glides and diphthongs for their identification and type categorization. A diphthong is the 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) [13, 4]. In order to quantify the acoustic characteristics of diphthongs and glides present in Sindhi; the main acoustic parameters measured are the first two formants (F1, F2), and the duration of three analysis points (on-glide, off-glide and the transition segment). 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 cues for the perception of diphthongs [14, 15, 16]. The speech samples obtained in two ways are gathered and the signals of the diphthong and glide segments were extracted from the word utterances and analyzed. The segment boundaries for the diphthong and glide phonemes in the word utterances were marked manually by inspecting the spectrograms of the word utterances. A set of acoustic properties for each diphthong phoneme of Sindhi were measured based on the speech signals of the twenty eight raw word utterances given in [17]. The formant frequencies for the first two formants (F1, and F2) were obtained using the Burg LPC (linear predictive coding) algorithm. The presence of the two vocalic elements in diphthongs and glides are determined by computing the z-scores for the first two formants according to the method described by authors [18, 19]. The z-scores for first two

formants of each diphthong and two glides 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 first two formants is to determine whether there are two vocalic elements corresponding to the diphthong phoneme or the speech signals belong to glide or other non-diphthong phoneme. In this paper the obtained negative (-ve) z-scores belong to one vocalic element and the positive (+ve) z-scores belong to the other vocalic element. 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 7 and 8 below show the computed z-scores for two glides of Sindhi and figures 9 to figure 10 show the z-scores for the two diphthong phonemes of Sindhi.

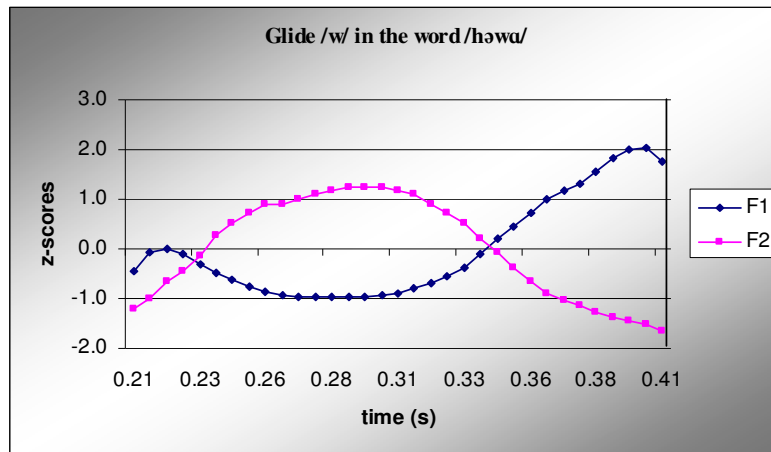


Figure 7. Typical first two formants z-scores of the glide /w/ present in the word /həwɑ/ هَوا of Sindhi. Note that the z-scores for both F1 and F2 are irregular.

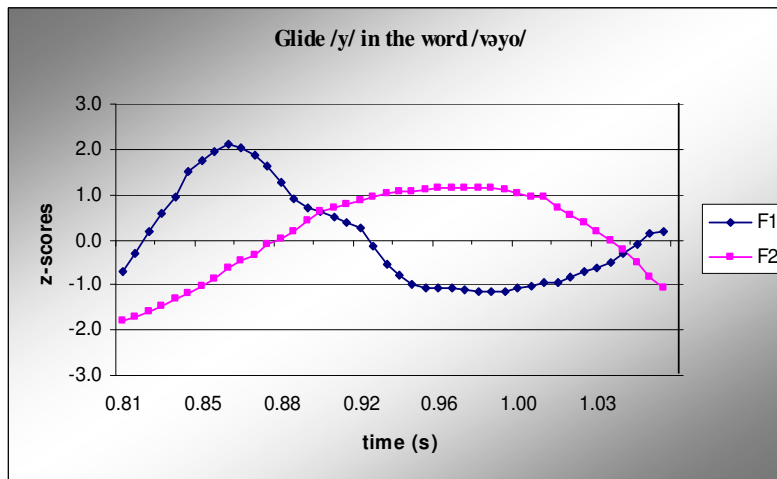


Figure 8. Typical first two formants z-scores of the glide /y/ present in the word /vəyo/ ويو of Sindhi. Note that the z-scores for both F1 and F2 are irregular.

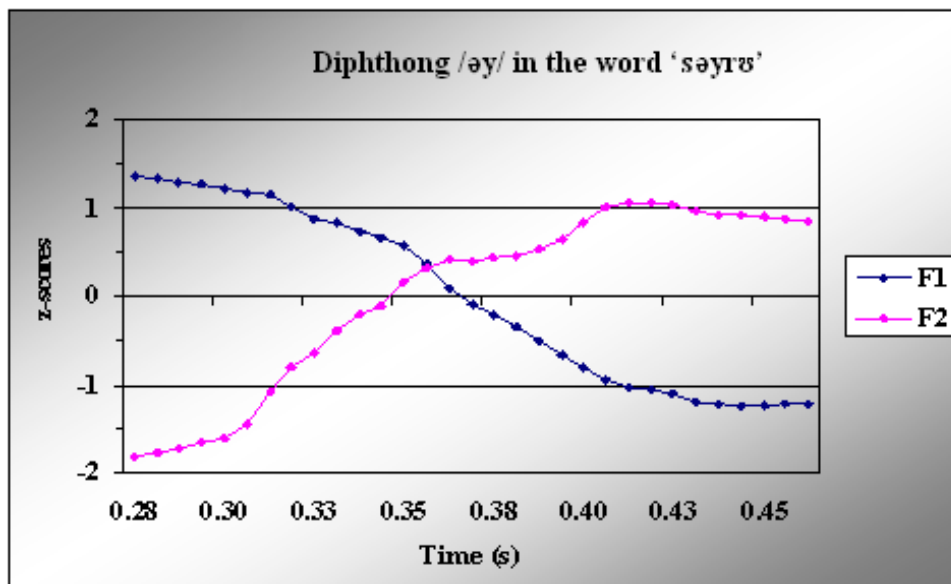


Figure 9. Typical first two formants z-scores of the diphthong [əy] present in the word utterance 's[əy]rə', of Sindhi. Note that F1 starts with +ve z-scores and moves gradually to the -ve z-scores and in similar way the F2 z-scores move gradually from -ve to +ve z-scores.

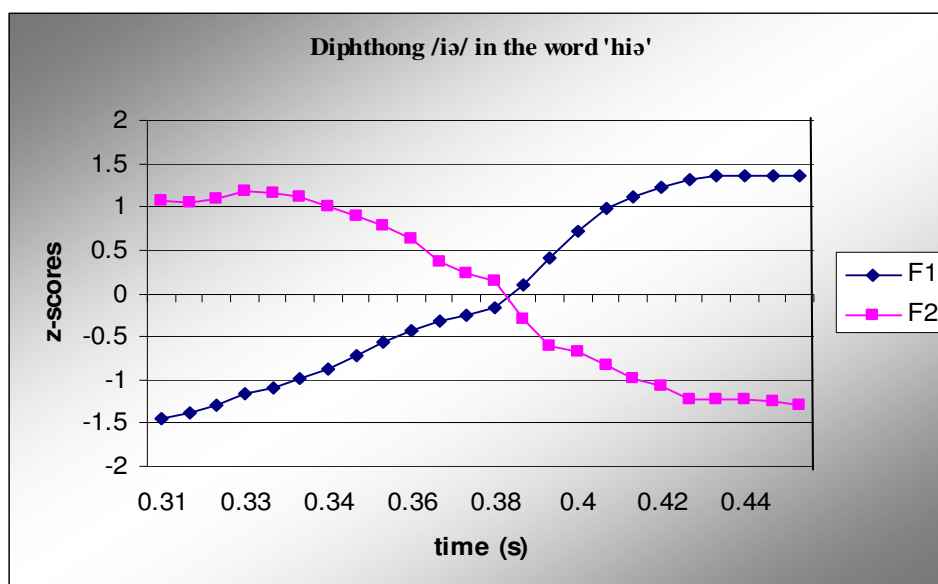


Figure 10. Typical first two formants z-scores of the diphthong [iə] present in the word utterance 'h[iə]', of Sindhi. Note that the F1 starts with -ve z-scores and move gradually to the +ve z-scores and F2 start with +ve z-scores and gradually move to the -ve z-scores.

Figure 9 shows 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. Figure 10 show that the diphthongs with the transition from close front vowel (first vocalic



element) to the open mid 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 F2 z-scores are +ve for first vocalic element and -ve for the second vocalic element. Typical z-scores for two diphthongs in figures 9 and 10 show that the F1 and F2 z-scores are stable at the beginning which is the duration of first vocalic element and then the transition period which shows a crossover of formant frequencies along with the gradual change in the formants motion followed by stability of the z-scores at the end which is the duration of second vocalic element. A significant difference among the z-scores of the diphthongs in Sindhi is 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. Shown in figures 7 and 8 are the typical z-scores for two glide consonants of Sindhi. Both the z-scores begin with -ve z-scores and two crossovers for both F1 and F2. The transition segment did not show a gradual formant motion rather a immediate sign change in the z-scores the first crossover sharply followed by another z-scores sign change the second crossover. The z-scores of two glides are not stable at the start neither they are stable at the end of glide consonant this ensures the absence of two vocalic elements (on-glide and off-glide) present in diphthongs of Sindhi. In other words we can conclude from the shown z-scores of figures 7, 8, 9 and 10 that more than one F1 and F2 z-score crossovers prove that the phoneme is glide consonant and one crossover ensures a diphthong phoneme of Sindhi.

Diphthongs are considered as moving voicing elements from the beginning vowel to the target vowel [6, 13]; whereas the glides are the sounds considered with a gradually changing formant pattern in the same way the pattern changes for the diphthongs [3, 7]. In this study the acoustic F1xF2 vowel plot of figure 11 shows how formants gradually move for the two vocalic elements of diphthongs and in which direction. The solid and dashed line arrows in figure 11 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 diphthongs [17]. The formant movement for diphthongs /aɪ/, /əy/, and /oɪ/ of Sindhi typically starts with the open vowel and end with the close-front vowel shown in figure 11. The formant movement for diphthongs [ɑʊ] and [əw] of Sindhi typically starts with the open vowel and end with the close-back vowel shown in figure 11. The formant movement of the diphthongs [iʊ] and [iə] of Sindhi typically starts with the close-front and end with the open-back vowel shown in figure 11. Acoustic vowel plot of figure 12 shows the formant movement for the glides /w/ and /y/ of Sindhi. As it is obvious from figure 11 that the diphthongs formants frequencies change from one vowel (on-glide) to other vowel (off-glide). The formants frequencies for the glide consonant do not show the moving formant pattern from one sound to the other sound this is shown in figure 12. The formants for glide /y/ show greater movement for the second formant whereas movement for the first formant is very small. The formants for glide /w/ showed very small movement for both (F1, F2) formants. Therefore it can be easily concluded that the two vocalic elements cannot be witnessed in the acoustic plot of figure 12; whereas it is obvious from the figure 11 that each diphthong phoneme showed formants movement between two vocalic elements.

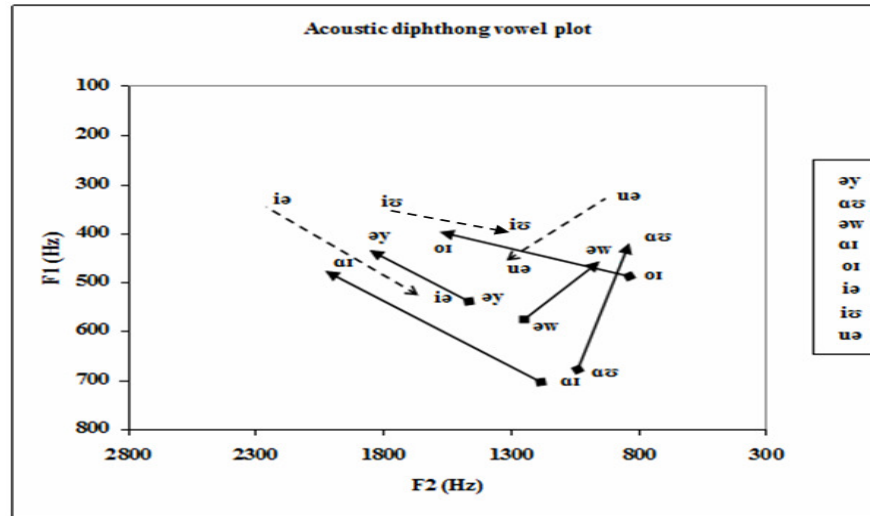


Figure 11. Acoustic F1 x F2 diphthong vowel plot of Sindhi [15]

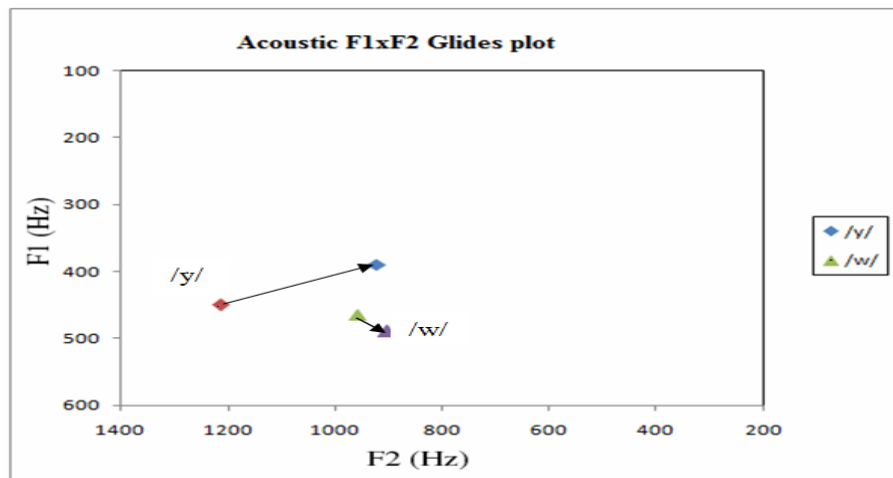


Figure 12. Acoustic F1 x F2 Glides plot of Sindhi.

## 5. CONCLUSION

Three types of Sindhi sounds (vowels, diphthongs and glides) are analysed and compared acoustically in this paper. The experiments based on the captured voice samples were conducted to analyse the steady state formant pattern for vowels and three critical analysis points for diphthong and glide phonemes of Sindhi. Since the diphthongs and glides are considered sounds with moving formant pattern, the formant transitions for these sounds were analysed at on-glide, off-glide and the transition segment of speech signals. The two vocalic elements present in diphthong phonemes are determined by analysing the sign transition based on computed z-scores for the first two formants. The acoustic F1xF2 vowel plot for the diphthongs is used to determine the formants movement from beginning vocalic element to the ending vocalic element. It is observed from the computed z-scores for first two formants that the glide consonants have more than one crossovers (sign transitions) and hence don't form a diphthong phoneme in Sindhi, regardless of its position in utterance; whereas the computed z-scores for diphthong phonemes

show one sign transition (crossover) for the computed z-scores of first two formants. Analysis results show that the transition segment for the diphthong phonemes is slow and more gradual compared to the glide consonants. The transition segment for glides is shorter in duration and form sharp transitory speech segments; because the vocal cords for the formation of glides move in and out quickly. It is also observed that the first two formants are sufficient to differentiate vowel and diphthong phonemes; whereas the second formant plays significant role to differentiate glides from the vowels and diphthongs.

## REFERENCES

- [1] Olive, J. P., Alice, G., & John. C. (1993). *Acoustics of American English Speech: a dynamic approach*. New York: Springer-Verlag.
- [2] Jennifer, S. C. (2006). The Sindhi language. In K. Brown (ed.) *Encyclopedia of Language and Linguistics* (2 ed., Vol. 11, pp. 384-386). Oxford: Elsevier.
- [3] Ladefoged, P. (1993). *A course in phonetics* (3 ed.). Harcourt College Publishers, New York.
- [4] MAXWELL, O. and FLETCHER, J. (2010), The acoustic characteristics of diphthongs in Indian English. *World Englishes*, 29: 27-44.
- [5] Kent, R. D., & Charles, R. (2002). *The acoustic analysis of speech* (2 ed.). Singular Publishing Group.
- [6] Raphael, L. J., Gloria, J. B., & Katherine, S. H. (2006). *Speech science primer: physiology, acoustics, and perception of speech* (5 ed.). Lippincott Williams & Wilkins.
- [7] Kehoe, M., G. Hilaire-Debove, K. Demuth & C. Lleó (2008) The structure of branching onsets and rising diphthongs: Evidence from the acquisition of French and Spanish. *Language Acquisition* 15: 5-57
- [8] Jones, D. (1969). *An outline of English phonetics* (9 ed.). England: W. Heffer & Sons Ltd., Cambridge.
- [9] Ioana, C. (2002). A perception-production study of Romanian diphthongs and glide-vowel sequences. *Journal of the International Phonetic Association* , Vol. 32, pp. 203-222.
- [10] Martínez Celdrán, E. (2004). Problems in the classification of approximants. *Journal of the International Phonetic Association*, Vol. 34, pp. 201-210.
- [11] Padgett, J. (2008). Glides, vowels, and features. *Lingua* , Vol. 118 (12), pp. 1937-1955.
- [12] Aguilar, L. (1999). Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech Communication* , Vol. 28 (1), pp. 57-74.
- [13] Gay, T. (1968). Effects of speaking rate on diphthong formant movements. *Journal of the Acoustics Society of America* ,Vol. 44, pp. 1570-1573.
- [14] Aiza, S., Sana, A., & Aymen, A. (2004). Diphthongs in Urdu Language and Analysis of their Acoustic Properties. *Center for Research in Urdu Language Processing (CRULP)*, pp. 9-15.
- [15] Borzone de Manrique, A.M (1979). Acoustic analysis of the Spanish diphthongs. *Phonetica*, Vol. 36, pp. 194-206.
- [16] Jha, S. K. (1985). Acoustic analysis of the Maithili diphthongs. *Journal of Phonetics*, Vol. 13, pp. 107-115.
- [17] Keerio, A., Patoli, M. Z., Mitra, B. K., Chatwin, C., Young, R., & Birch, P. (2010). Acoustic Analysis of Diphthongs in Sindhi. *GRASSROOTS* , 41 (1), pp. 1-18.
- [18] Hongyan, W. (2007). Mutual intelligibility of Chinese, Dutch and American speakers of English. Ph.D dissertation, Graduate School of Linguistics: Netherlands.
- [19] Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America* , Vol. 49 (2B), pp. 606-608.

### Authors

**Dr. Ayaz Keerio** is an assistant Professor at the Institute of Mathematics and Computer Science (IMCS), University of Sindh, Jamshoro, Pakistan. He got his Master's degree in Computer Science from University of Sindh, Jamshoro (Pakistan) and PhD from University of Sussex, UK in 2011. His main area of research is Speech Recognition and Synthesis systems. He is also interested in digital signal processing, Data communication & networks and mobile & distributed computing systems.



**Dr. Lachhman Das Dhomeja** is an Assistant Professor at the Institute of Information & Communication Technology (IICT), University of Sindh, Jamshoro, Pakistan. He got his Master's degree in Computer Technology from University of Sindh, Jamshoro (Pakistan) in 1991 and PhD from University of Sussex, UK in 2011. His main research area is Pervasive Computing in general and policy-based context-awareness in particular. His other research interests include secure device pairing in ubiquitous environments, Data communication & networks, software architectures and Distributed Computing.



**Dr. Asad Ali Shaikh** is an Associate Professor and director of the Institute of Information and Communication Technology (IICT), University of Sindh, Jamshoro, Pakistan. He did his Masters degree in Computers Engineering from Clarkson University, USA in 1991 and PhD degree in Information Technology from University of Sindh, Pakistan in 2006. His current research focus is on the protocol design, security issues in computer networks and software development. He is also interested in digital signal processing and Data communication & networks.



**Dr. Yasir Arfat Malkani** is a Lecturer at the Institute of Mathematics and Computer Science (IMCS), University of Sindh, Jamshoro, Pakistan. He got his Master's degree in Computer Science from University of Sindh, Jamshoro (Pakistan) in 2003 and PhD from University of Sussex, Brighton, UK in 2011. His main area of research is Pervasive Computing. His research is focused on secure device/service discovery and access control mechanisms using policies and location/proximity data/information. He is also interested in sensor networks, wireless networks (including WiFi, Bluetooth, WiMAX, etc), and solutions to various issues in distributed and pervasive computing systems through the integration of tools and techniques from distinct disciplines/areas. He is also interested in the design and/or development of various tools and techniques that might be useful in giving world-wide recognition to various national languages, such as SINDH and URDU.

