



Acoustics of Isolated Vowel Sounds of Sindhi

A. KEERIO⁺⁺, N. CHANNA^{**}, B. MITRA, R. YOUNG, C. CHATWIN

Department of Computer Science, University of Sussex, United Kingdom

Received 27th December 2013 and Revised 19th May 2014

Abstract: Sindhi is widely spoken in the Sindh province of Pakistan, and in some parts of India. This paper presents the acoustic analysis of the isolated utterances of all ten vowels present in the language inventory. The entire study is based on 7,500 voice samples collected through a rigorous field study from 75 native male speakers of the Sindh region in Pakistan. Subtle and major differences of the vowel qualities of the five distinct dialects of the region are highlighted. Finally the general acoustic features of the vowels considering the region as whole are described

Keywords: Formants, Frequency, Fundamental frequency, acoustic, vowel qualities.

1. INTRODUCTION

Much research has been done in the past on developing techniques to study the acoustic phonetics of the isolated sounds of the elements of a language-set, and in relating those to the other branches of phonetics (like articulatory). The impetus for doing such research comes from the fact that detailed knowledge of the sound characteristics of the language-elements has a broad variety of applications - from developing accurate synthetic speech production systems to modelling robust speaker-independent speech recognizers (Ansarin 2004, Lee 1992). It has also been noticed that vowel-sounds are studied more in comparison to consonant-sounds as the former show a steady state acoustic pattern in the speech stream (Kent 2002, Moore 2003). On the other hand, the language whose elements in the inventory have been studied the most is either English or foreign-accented English (e.g., Peterson 1952, Chen 2001, Maxwell 2009, Deterding 2003). Other languages whose vowel sounds have also been thoroughly studied are Persian (Ansarin 2004), French (Ouni 2003), Shanghai Chinese (Chen 2008) and Turkish (Gordon 2006), etc. However, to the best of our knowledge, no such comprehensive work has ever been undertaken to study the acoustics of the vowel-sounds of Sindhi. Work reported in the literature to date on the language has either been on articulatory phonetics of the element-sounds, or on the language’s writing system, grammar, dialects and history. In this respect, note that Raza *et. al.*’s work reported in their Centre for Research in Urdu Language Processing (CRULP) Annual Report 2003-2004 (Raza 2009) has been on articulatory phonetics of the language’s consonantal inventory and on acoustic phonetics of the voiced implosives; the authors have thoroughly compared their findings with those of Urdu. Moreover, the report is based on samples collected from

the native speakers of only one district (Jacobabad) in Sindh, and hence does not include the diversity of the five distinct dialects of the province. Note here, there are five different regional dialects present in the Sindh province in Pakistan; the only other distinguishable dialect is spoken in the Rann of Kutch—a region in the western state of Gujarat, India (Jennifer 2006). Jatoi’s work, written in Sindhi, (Jatoui 1996) in this field has been on the articulatory phonetics of the vowel- and consonant- sounds of the language, and also on the history of the language. It should also be noted that the vowel plots reported in the CRULP report and in Jatoui’s book differ in the vowel positions (and, hence, in vowel qualities) significantly for the three following vowels: (i) /ə/ (ii) /a/ and (iii) /o/. As a part of this research work, a rigorous field study was undertaken to collect voice samples from the native speakers of Sindh, Pakistan; all the five distinct dialects have been equally represented in the data set so that subtle and major differences in the isolated vowel-sound characteristics can be studied. The paper presents a thorough acoustic analysis of the isolated vowel sounds collected from native male speakers of the five distinct regions of the province. The exhaustive analysis results reported in the paper will not only help us in gaining further insights about the language’s elemental vowel sounds but also would help in designing an artificial Sindhi speech production system or a robust Sindhi speech recognizer.

The paper is organized as follows: section 2 describes the methodology and details of the field study conducted; section 3 discusses the signal pre-processing method used in the current work; section 4 talks about the IPA vowel diagram and acoustic features of the Sindhi vowels; section 5 presents results and discussions and finally the conclusions are drawn in section 6.

⁺⁺Correspondence to: Email: ayazkeerio@hotmail.com, nchanna68@yahoo.com, Cell. No. +92-3003069048

^{*}Institute of Mathematics & Computer Science, University of Sindh, Jamshoro

^{**}Institute of Business Administration, University of Sindh, Jamshoro

2. MATERIAL METHODS

The geographical region selected for the study is the Sindh province of Pakistan. This region includes five language dialects out of the six dialects of Sindhi outlined in (Jennifer 2006). The only language dialect not included in the study is spoken in the Rann of Kutch located in the western part of Indian state Gujarat. Seventy five adult male speakers participated in the study (fifteen speakers for each dialect). The selection of the speakers was done after an informal interview with each speaker to judge whether the speaker has any speaking disorder.

2.1 Recording environment

Two different types of recording environments are observed in the field study: (i) open environment and (ii) closed environment. Recordings made while sitting with a participant alone in an open-air places called “jhupri’s or *Otaks*” are treated as open environment, whereas recordings made while sitting with a participant inside a concrete building are treated as closed environment.

2.2 Sample Recording

For the recording of voice samples each speaker was assigned a sheet containing the ten vowels written in Sindhi script. At least ten instances of each speaker’s voice samples were recorded for the purpose of enabling the best utterance selection later on. Tool used for voice sample recording include a portable solid state digital audio recorder (Marantz PMD660) used with an external broadcast quality condenser microphone (the RODE NTG-2). The speech samples were recorded as single channel (mono mode) at the sampling rate of 48 kHz; the output of the channel was obtained as 16 bit words. The recorded voice samples were stored as uncompressed .wav files for subsequent offline processing. The first region selected for the recording of voice samples was the place where the *Middle* dialect of Sindhi is spoken. The Hyderabad district was set as the origin to calculate the approximate distances travelled to cover all the five dialects of the province. Approximate road distance to various locations from the origin in kilometres, and the number of speaker’s voice samples taken from one single location are shown in (Table 1).

Table 1. Approximate road distance to other locations from origin (Hyderabad)

Dialect	Location	Distance from Origin - km	Number of speaker’s participated for recording
	Hyderabad	--	04
Middle	Jamshoro	18	04
	Sakrand	107	03
	Nawabshah	123	04

Utradi	Khairpur	305	04
	Pano Aqil	357	04
	Shikarpur	376	03
	Larkana	470	04
Lari	Badin	110	04
	Talhar	87	04
	Thatta	102	04
	Sujawal	125	03
Thareli	Mithi	150	07
	Islamkot	240	08
Lasi	Thano Bula Khan	72	08
	Thano Ahmed Khan	85	07

3. SIGNAL PRE-PROCESSING

Pre-processing is required for several reasons in the speech processing systems i.e. to remove ambient noise signals or to segregate the voiced part from the unvoiced/silence parts of the speech signals (Keerio 2009). The noise free speech signals help to determine various acoustic parameters accurately and to improve the overall system performance for subsequent offline analysis process. There are several signal pre-processing methods available; the most commonly used methods are: (i) the short time energy (STE) based methods, (ii) the distribution based methods, and (iii) the statistical outlier detection based methods (Keerio 2009). In this work STE based method of signal pre-processing is used; however the statistical outlier detection based method can also be used with certain constraints on the data set as discussed in (Keerio 2009).

3.1 The STE based non-overlapping moving window method

STE based methods of speech signal segregation are usually computationally efficient. These methods utilize the fact that the energy in the voiced region is greater than in the silence/unvoiced region (Atal 1976, Mitra 2005, Keerio 2009). In this study the voice samples of the recorded utterances were partitioned into windows of 5 ms duration. The mean for each 5 ms window is than computed and plotted against a time scale shown in **Fig. 1(b)**. The threshold for the selection or rejection of the N^{th} 5ms window was computed as the mean of the two 5 ms windows from the voiced region samples. The first 5ms window that shows a peak (which usually marks the start of the voice samples) is chosen and a second 5 ms window with equal or bit lower in peak is chosen. One typical example is shown in **Fig. 1(b)**: the first 5 ms window starts at approximately 130 ms (first vertical line in Figure 1(b)) and the second 5 ms window starts at approximately 500 ms (second vertical line in Figure 1(b)). If the mean of the N^{th} 5 ms window is more than the threshold, then the samples of the window were retrieved as voiced samples, otherwise they were dropped.

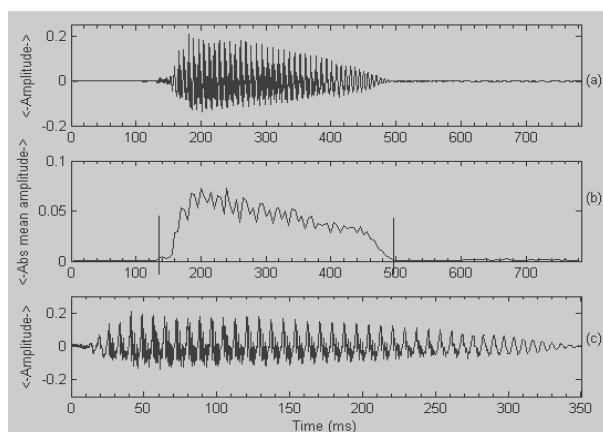


Fig. 1. Typical waveform of a vowel utterance: (a) speech samples before the pre-processing method was applied (b) means of 5ms windows (c) speech samples after the pre-processing method was applied.

4. IPA vowel diagram and Sindhi vowel system

Vowels are generally difficult to define in terms of articulatory phonetics because the main articulatory organs involved in changing the shape of vocal tract for the passage of air during the production of the vowels are the tongue and the lips (Jones 1969). During the formation of vowel sounds, the tongue moves towards the other articulators; for example, the body of the tongue is raised up in the direction of the hard palate for high front vowels and in the direction of the soft palate for high back vowels. The formation of Sindhi vowel sounds can also be best described in terms of these three main articulators in the oral cavity (Jatoi 1996): (i) the tongue (ii) the lips and (iii) the mouth. However due to the lack of resources required to quantify the articulatory properties of the Sindhi vowels physiologically (such as X-ray imaging, Magnetic Resonance Imaging, Electromagnetic Midsagittal Articulography etc) this study rely on the acoustic analysis of Sindhi vowels. The articulatory descriptions in this study are based on the material available for the English language on the subject of articulatory phonetics.

4.1 Acoustic features of Sindhi vowels

A set of acoustic properties for each vowel were measured across all five dialects (15 speaker samples per dialect adding up to 75 speaker voice samples in total). The main acoustic parameters measured are: the mean fundamental frequency (F0), the mean of the first four formants (F1, F2, F3, and F4), and the mean vowel duration. The formant frequencies for the first four formants were obtained using the Burg

linear predictive coding (LPC) algorithm implemented in Praat speech processing tool (Boersma 2013).

4.2 Vowel duration

The phonemic inventory of Sindhi vowels comprises the set of ten vowels given in Table 2. Phonologically the Sindhi vowel system can be divided into two vowel classes according to the vowel length contrast: (i) short (lax) vowels and (ii) long (tense) vowels. In Sindhi vowel system, the vowel durational differences have been observed for three long and short vowel pairs; these pairs are: (i) /i/ and /ɪ/ (ii) /u/ and /ʊ/ and (iii) /ɑ/ and /ɔ/. The vowel duration in this study data set has been measured in two different ways across the five regional dialects. In first method the vowel samples were trimmed (by excluding the leading and trailing silence or background noise regions) and then the vowel duration calculated. In second method the spectrograms for the vowel utterances generated by using the Praat Speech processing tool (Boersma 2013) and the vowel start and end positions were marked through the visual inspection of the spectrograms and then the vowel duration calculated. The two obtained durations for each vowel are compared, and the mean of the two durations is assigned to a target Sindhi vowel. It is observed that speakers of the same geographical region produce different vowel durations for the same vowel, because of the fact that some speakers may speak fast and some may speak slow (Rosner 1994). In this situation it is barely possible to conclude whether the vowel lies either in the class of long (tense) or short vowels of the inventory from the computed raw vowel durations. In this respect the vowel normalization procedure is adopted to determine whether a vowel is a long or a short vowel as described by (Hongyan 2007, Lobanov 1971). In the normalization procedure the z-scores for the entire Sindhi vowel set were computed by subtracting the mean vowel duration from the raw vowel durations. The differences are then divided by the standard deviation of the entire vowel set. The mean of the obtained z-scores for each vowel is then computed; a negative z-score for a vowel categorizes the vowel as a short vowel, and a positive z-score categorizes the vowel as a long vowel. The vowels /ɪ/, /ʊ/, and /ɔ/ of the Sindhi vowel system produced negative z-scores, and are classified as short vowels; the vowels /i/, /u/, /ɑ/, /æ/, /ɔ/, /o/, and /e/ produced positive z-scored means, and are classified as long vowels of the inventory; the mean vowel durations for ten Sindhi vowels are shown in (Table 2).

Table 2. Vowels in the inventory.

IPA symbol	/i/	/ɪ/	/e/	/ɛ/	/ə/	/ɑ/	/ɔ/	/o/	/ʊ/	/u/
Sindhi Symbol	ای	اِ	ای	ای	ا	ا	او	او	ا	او
Duration in ms	344	169	370	358	194	334	304	390	181	367

5. RESULTS AND DISCUSSION

Formants provide the fundamental cues to analyze and quantify the perceived vowel quality measures. Formants are usually present in all the voiced phonemes of a language such as vowels, nasals, and approximants etc; however in vowels they are strong and relatively stable over time (Moore 2003). For acoustic phonetic analysis the first two formants are sufficient; however for recognition and synthesis three to five formants are required (Parsons 1987). Formants may vary for the same vowel between speakers of different age and sex; they may also vary in more than one instance of the same speaker for a given vowel utterance; however that variation may produce a subtle difference in the vowel quality. This variation is due to the fact that the vocal tract is a tube of non-uniform cross sections, approximately 17 cm in length for adult males (Parsons 1987). The frequencies at which it resonates naturally occur at odd multiples of 500 Hz (Raphael 2006, Parsons 1987). Since the cross section is non-uniform, the vocal tract resonant frequencies are not equally spaced (Parsons 1987). Thus, it is hardly possible to associate a constant frequency value to these resonances known as formants. However the formants can be defined in a range of frequencies; F1 ranges between 200-800Hz, F2 ranges between 600-2800Hz, and F3 ranges between 1300-3400Hz for an adult male speaker (Parsons 1987). Phoneticians agree that the vowel quality can be quantified acoustically by measuring the lower two resonant frequencies (F1 and F2) of the vocal tract. Therefore in this study acoustic analysis of the Sindhi vowel system relies on the first

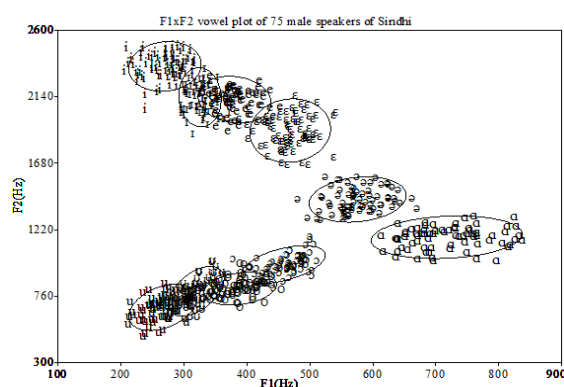
two formants. A variety of published literature is available on the subject of vowel quality and intelligibility based on the analysis of the first two formants. The two basic articulatory characteristics that roughly correspond to the first two formants are: (i) the tongue height and backness and (ii) the lip rounding (Pfitzinger 2003, Ladefoged 1993). The first formant is closely correlated to the perceived vowel height (Raphael 2006) in a way that lower the F1 value the higher is the tongue body in the mouth, and the converse is also true. Similarly, the second formant is closely related to the perceived tongue backness and the lip rounding (Raphael 2006) in a way that the lower the F2 value the more backward is the tongue body in the mouth, and the lips are rounded; conversely the higher the F2 value the more forward is the tongue body in the mouth and the lips remain either neutral or widely spread. There are many languages for which acoustic vowel plots have been drawn based on the first two formants and comparative conclusions are made for their vowel system. To the best of our knowledge no such acoustic phonetic vowel plot based on the first two formants for Sindhi across five language dialects is available in the literature. The purpose of the field study is that we have conducted and the acoustic phonetic analysis carried out on the captured voice samples provides this knowledge and data for the research community. Shown in Table 3 below are the mean formant frequencies of the first four formants, the mean fundamental frequency, the mean duration and the standard deviation of ten Sindhi vowels across the five dialects of the Sindh province Pakistan.

Table 3. The mean F0, F1, F2, F3 and F4 frequency values in Hz, and the duration and the standard deviation obtained from the voice samples of 75 adult male speakers of Sindhi.

Dialects		Middle		Utradi		Thareli		Lari		Lasi	
Vowel	Parameter	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
/i/	F0	160	-	179	-	166	-	166	-	160	-
	F1	256	24	241	28	261	30	282	19	268	19
	F2	2330	49	2331	85	2341	106	2342	96	2409	62
	F3	3033	239	3103	164	3005	271	3110	229	3113	183
	F4	3640	267	3741	175	3652	310	3752	176	3732	139
	Duration	339	24	328	32	360	22	340	32	357	32
/ɪ/	F0	163	-	167	-	132	-	160	-	128	-
	F1	324	17	318	25	327	18	325	19	331	21
	F2	2153	70	2084	92	2038	115	2120	90	2185	147
	F3	2752	250	2600	242	2720	206	2714	149	2754	209
	F4	3669	309	3694	284	3788	324	3765	169	3757	176
	Duration	167	25	159	20	175	23	175	22	171	45
/e/	F0	154	-	144	-	135	-	161	-	160	-
	F1	402	28	383	21	371	22	380	30	366	32
	F2	2042	67	2031	106	2077	106	2105	86	2155	80
	F3	2727	209	2725	186	2739	155	2730	133	2812	218
	F4	3612	269	3704	316	3682	195	3739	101	3583	242
	Duration	375	28	379	34	366	23	369	42	362	22
/ɛ/	F0	130	-	147	-	140	-	155	-	156	-
	F1	472	26	451	37	464	32	492	31	463	22
	F2	1949	76	1910	99	1945	103	1952	106	1958	122
	F3	2551	353	2593	143	2645	174	2619	122	2597	303

	F4	3421	468	3691	211	3641	209	3661	354	3408	486
	Duration	365	30	370	22	346	18	353	43	358	24
/ə/	F0	147	-	145	-	137	-	145	-	157	-
	F1	594	28	577	47	587	37	553	42	567	24
	F2	1490	58	1386	77	1437	66	1419	82	1413	76
	F3	2533	326	2667	230	2700	194	2604	166	2632	197
	F4	3582	400	3735	121	3700	122	3693	199	3591	314
	Duration	215	20	232	29	195	22	172	23	160	15
/a/	F0	132	-	170	-	151	-	140	-	145	-
	F1	755	56	700	47	714	57	714	63	700	62
	F2	1215	55	1148	78	1147	76	1148	73	1181	68
	F3	2540	285	2574	337	2681	105	2566	210	2588	181
	F4	3570	410	3450	422	3758	218	3529	323	3624	137
	Duration	330	28	352	24	318	17	326	22	343	32
/ɔ/	F0	136	-	123	-	152	-	159	-	141	-
	F1	447	28	449	32	462	37	460	34	469	35
	F2	928	45	952	83	1016	62	981	33	1019	44
	F3	2473	190	2505	190	2620	110	2532	111	2566	260
	F4	3505	331	3597	311	3548	228	3679	112	3577	138
	Duration	295	34	312	41	305	31	285	38	325	21
/o/	F0	163	-	151	-	170	-	145	-	148	-
	F1	386	28	375	25	387	21	419	18	377	29
	F2	814	36	840	53	795	74	832	31	825	73
	F3	2481	340	2509	459	2598	182	2597	150	2668	129
	F4	3437	291	3420	402	3566	170	3368	229	3543	277
	Duration	388	23	380	41	379	39	407	30	398	38
/ɔ:/	F0	162	-	168	-	157	-	159	-	179	-
	F1	316	14	328	28	324	20	327	23	303	15
	F2	773	57	870	84	764	99	813	70	750	65
	F3	2618	130	2490	128	2470	255	2452	143	2577	159
	F4	3580	131	3504	178	3486	275	3432	262	3647	239
	Duration	181	22	180	21	171	17	189	23	187	29
/u/	F0	165	-	129	-	165	-	162	-	165	-
	F1	279	24	252	26	269	29	286	21	283	23
	F2	679	57	645	82	714	83	731	59	645	84
	F3	2454	245	2391	372	2540	205	2551	147	2340	407
	F4	3518	236	3416	341	3580	331	3485	230	3347	422
	Duration	375	37	357	38	351	41	385	42	369	22

The (Fig. 2) below show the F1 x F2 scatter vowel plots for the utterance of the seventy five adult male speakers of Sindhi. The ellipses fitted to the data are drawn for each vowel category to cover one vowel data on the vowel plot. It is clear from the vowel plot (Fig. 2) that the vowel positions are relatively stationary across all five dialects of Sindhi. However, due to the different regional dialects a greater variation is observed in the F1 value for the low back (open back) vowel /a/ and the mid open central vowel /ɔ/ throughout the five dialects. The high front and back vowel pairs: /i/-/ɪ/ and /u/-/ʊ/ are observed varying in the F2 frequency value through out all five dialects. Tight vowel clusters are also evident in (Fig. 2) across all five dialects; moreover the overlapping in the vowel space between the high front vowel pair /i/ and /ɪ/ and the high back vowel pair /u/ and /ʊ/ is observed. The neighboring front and back series vowels /ɪ/, /e/, /ɔ/, /o/, /ɛ/ and /ɔ:/ overlap for all five dialects as shown in fig. 2. The open back vowel /a/ and mid open vowel /ɔ/ show clear vowel quality across all five dialects shown in (Fig. 2).



(Fig. 2). The F1 x F2 scatter vowel plot for 75 adult male speakers of Sindhi language. Ellipses are drawn fitting the vowel space for each vowel.

The association of the corresponding vowel quality and the formant frequencies is non-linear. So the formant values are converted using the Bark transform formula (Traunmuller 1990, Hongyan 2007). The

distance between two vowel qualities are then computed using the F1 and F2 values (in Bark). Fig.3 below show the Bark transformed acoustic phonetic vowel plot based on F1 and F2 values for the Sindhi vowel system. In fig. 3 the F1 values (in Bark) are plotted on the ordinate (vertical axis) and the F2 values (in Bark) are plotted on the abscissa (horizontal axis); the axis values are inverted so that the maximum crosses at the origin. The four corner vowels of the Sindhi vowel system can then be selected with the help of the acoustic phonetic vowel plot of figure 3. By drawing a line between the four corner vowels in figures 3 starting from the high front vowel /i/ to /ɛ/, from /ɛ/ to /a/, from /a/ to /u/ and finally from /u/ to back /i/, we get a four sided figure; which covers all the remaining vowels of the Sindhi vowel system. The four sided figure has roughly a similar shape to the IPA vowel quadrilateral. The vowels marked by this four sided figure are referred to as the extreme reference points or corner vowels of Sindhi. These four corner vowels of Sindhi /i/, /ɛ/, /a/, and /u/ shown in (Fig. 3) can be used as fixed reference points for the classification and articulatory description of the other six vowels of the inventory according to the eight cardinal vowels of English as described by (Jones 1969) and the four extreme reference English vowels described in the IPA hand book (IPA 1999). The midpoint calculated in the acoustic vowel plot in bark unit help to classify the vowels of Sindhi inventory as front and back series vowels also the high front or low front and high back or low back vowels of the Sindhi vowel system.

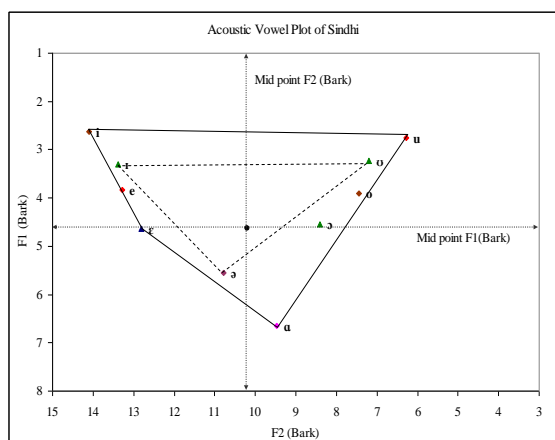


Fig.3. Acoustic vowel diagram of Sindhi vowel system showing the mean formant frequencies (in Bark) for seventy five native male speakers of Sindhi the axis values are inverted. The solid line connects the four corner vowels and the dashed line connects the three short vowels of the Sindhi vowel system.

6.

CONCLUSION

In this paper a thorough study has been conducted on the isolated vowel-sounds of Sindhi,

covering all the five dialects spoken in the Sindh province of Pakistan. The reported acoustic parameters of the sounds include the fundamental frequencies, the average durations, and the first four formant frequency values. Based on the average durations, the vowels present in the language inventory have been categorized either as lax or tense vowels. Using the first two formant frequency values and knowledge about the relationships between acoustic phonetics and articulatory phonetics, a clear picture can be derived about how the articulators will move while producing each of the ten vowel sounds of the language. From the vowel plots presented in the paper, an idea about the vocal exercise that needs to be performed to produce the ten isolated vowel-sounds and the intelligibility of each of the vowel-sounds can be formed. The classification of Sindhi vowels in terms of front and back series vowels and the understanding of high or low vowels of the Sindhi vowel system can be observed. In the future, the parameter values reported in this paper will help in developing dialect specific Sindhi vowel-sound synthesizers and digital recognition systems.

REFERENCES:

- Ansarin, A. A. (2004). An acoustic analysis of modern Persian vowels. In *The 9th SPECOM-2004 Conference, "Speech and Computer", Saint-Petersburg, Russia*, http://www.isca-speech.org/archive/specom_04 10th Jan.2014
- Atal, B. and L. Rabiner (1976) A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. (24): 3, 201-212.
- Boersma, P., and D. Weenink. (2013). Praat: doing phonetics by computer 5314 <http://www.praat.org/> (10 December 2013).
- Chen, Y., M. Robb, H. Gilbert, and J. Lerman, (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics*, vol. (15): 427-440.
- Chen, Y. (2008). The acoustic realization of vowels of Shanghai Chinese. *Journal of Phonetics*, Vol. (36): 4, 629-648.
- Deterding, D. (2003). An instrumental study of the monophthong vowels of Singapore English. *English World Wide*, vol. (24): 1, 1-16.
- Gordon, M. and A. Applebaum. (2006). Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association*, vol. (36): 2, 159-186.

- Hongyan, W. (2007). *Mutual intelligibility of Chinese, Dutch and American speakers of English*. Ph.D. dissertation. Netherlands Graduate School of Linguistics.
- IPA. (1999). *Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press., USA.
- Jatoi, A. 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, vol. (11): 384-386.
- Jones, D. (1969). *An outline of English phonetics*. 9th edn (reprint). England: W. Heffer & Sons Ltd., Cambridge. University Press., USA.
- Keerio, A., B. K. Mitra. P. Birch. R. Young. and C. Chatwin. (2009). On preprocessing of speech signals. *International Journal of Signal Processing*, vol. (5): 3, 216-222.
- Kent, R. D. and R. Charles. (2002). *The acoustic analysis of speech*, 2nd edn. Singular Publishing Group.
- Ladefoged, Peter. 1993. *A course in phonetics*, 3rd edn, Harcourt College Publishers, New York.
- Lee, C. H., E. Giachin, L. R. Rabiner, R. Pieraccini, and A. E. Rosenberg. (1992). Improved acoustic modelling for large vocabulary speech recognition. *Comput. Speech Lang.*, vol. (6): 103-127.
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, vol. (49): 2B, 606-608.
- Maxwell, O. and J. Fletcher. (2009). Acoustic and durational properties of Indian English vowels. *World Englishes*, vol. (28): 1, 52-69.
- Mitra. A., B. Chatterjee, B. K. Mitra. (2005). Identification of primitive speech signals using TMS320C54X DSP processor. In *The 11th national Conference on Communications (NCC), IIT-Kharagpur, India*.
- Moore, B. C. J. (2003). *An introduction to the psychology of hearing*, 5th edn. Emerald Group Publishing. China.
- Unni, S. and Y. Laprie. (2003). A study of the French vowels through the main constriction of the vocal tract using an acoustic-to-articulatory inversion method. In *The 15th International Congress of Phonetic Sciences, Barcelona, Spain, 2901-2904*.
- Parsons, T. W. (1987). *Voice and speech processing*, McGraw-Hill. Spain,
- Peterson, G. E. and H. L. Barney. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 175-184.
- Pfifzinger, H. R. (2003). Acoustic correlates of the IPA vowel diagram. In *The 15th International Congress of Phonetic Sciences, Barcelona, Spain, 1441-1444*.
- Raphael, L. J. and J. B. Gloria S. H. Katherine, (2006). *Speech science primer: physiology, acoustics, and perception of speech*, 5th ed, Lippincott Williams & Wilkins.
- Raza, S. and F. Z. Agha, (2009) Usman R. Phonemic inventory of Sindhi and acoustic analysis of voiced implosives. *Center for research in Urdu language processing*, <http://www.crupl.org/research/reports/streport04.htm> (12 Jan 2013)
- Rosner, B. S., and J.B Pickering. (1994). *Vowel perception and production*. Oxford: Oxford University Press. UK.
- Trautmuller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, (88): 1, 97-100.