Acoustic Characteristics of Pakistani English Vowel Sounds

Abdul M. Abbasi[1], Mansoor A. Channa[2], Masood A. Memon[3], Stephen John[4], Irtaza Ahmed[5] & Kamlesh[5]

[1] Department of Allied Subjects, Sindh Madressatul Islam University, Karachi, Pakistan

[2] Department of Basic Sciences, Quaid-e-Awam University, Nawab Shah Sindh, Pakistan

[3] Department of English, Cadet College Petaro, Hyderabad, Sindh, Pakistan

[4] Department of Education, Sindh Madressatul Islam University, Karachi, Pakistan

[5] Department of Computer Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

Correspondence: Abdul M. Abbasi, Department of Allied Subjects, Sindh Madressatul Islam University, Karachi, Pakistan. E-mail: amabbasi@smiu.edu.pk

Abstract

The purpose of this investigation was to document acoustic characteristics of Pakistani English (PaKE) vowel sounds. The experiment was designed to examine the properties of ten vowels produced by Pakistani ESL learners. The analysis is based on the voice samples of recorded 50 CVC words. Total 5000 (10× 10 × 50=5000) voiced samples were analyzed. The data consisted of 50 words of ten English vowel sounds [iː ɪ e ɔː æ ə ɑː uː ɒ ʊ]. Ten ESL speakers recorded their voice samples on Praat speech processing tool installed on laptop. Three parameters were considered i.e., fundamental frequency (F0), vowel quality (F1-F2) and duration. Formant patterns were judged manually by visual inspection on Praat *Speech Processing Tool*. Analysis of formant frequency shows numerous differences between male and female of F1 and F2, fundamental frequency and duration of English vowels. The voice samples provide evidence for higher and lower frequency of vowel sounds. Additionally, the data analysis illustrates that there were statistical differences in the values of short and long vowels coupled with vowel space plot showing explicit differences in locating the production of vowels of male & female vowel space acoustic realizations.

Keywords: acoustic characteristics, vowel sounds, frequencies, Pakistani ESL

1. Introduction

This study is based on 5000 voice samples (10× 10 × 50=5000) which were analyzed in a very careful manner. The experiment was designed to document English vowel properties produced by Pakistani English speakers. To analyze PaKE vowels, three phonetic exponents were considered as parameters for the analysis e.g., duration, fundamental frequency and first and second formant frequency. The study focuses on the characterization of vowel sounds produced by English as Second Language (ESL) learners and the positions of Pakistani vowels space in vowel space plot of mouth cavity. The objectives behind this study are to document the variations in frequencies of 10 different English vowels by ESL learners. Every vowel has different frequency and the vowel space plot when it is pronounced by different people it contains a different frequency, plot, and duration coupled with different pitch. Praat speech processing tool was utilized to analyze the sounds with the help of spectrographic view. Praat calculated the average of all 5000 sounds. The measurements were categorized in terms of male, female and third their measurements were compared with each other. The study tests the hypothesis that if ESL learners pronounce English vowels, then their placement of vowels on vowel space will be differently placed by male and female Pakistani ESL learners. Based on the research questions as follows: What are the acoustic realizations of PaKE English vowel sounds? What is difference in placement of acoustic vowel space plot between male and female?

Additionally, the study focuses on acoustic properties of English vowel sounds and their acoustic vowel space in vowel inventory plot. The vowels are basically categorized referencing three phonetic factors i.e., height of the vowel, back-ness and lip rounding. Olive, Greenwood, & Coleman (1993) contends "An acoustic analysis of vowels stresses the different formant configurations that are characteristic of each vowel. The relationship among the vowels can be examined by comparing their formant values". Front-back and high-low differences are

characterized by the second and the first formants, illustrating on spectrographic images. If the formant value is high, the closer the vowel becomes to the front point. In so far as front vowel is concerned frequency F1 turns out to be lower and contraction escalates in the vocal track. F1 and F2 frequencies are regarded as a vowel quality in acoustic speech terms. Generally unstressed vowels are reduced in quality; whereas the stressed vowels have higher frequencies. There are several features in the production of sounds, i.e., reduction of vowel quality, vowel lengthening followed by the voiced sound. Another important factor is absence and presence of lexical stress which affects the vowel quality. The unstressed syllables in American English are reduced in quality therefore, they involve in vowel centralization to schwa vowel sound, the chiefly produced sound in English language. F1 and F2 formants have greater distance when there occurs in between the highest vocal sounds. Whereas, for back vowel, F2 is much closer and lower to F1 for the back vowels than for the front vowels (Parson, 1987).

2. Methods

2.1 Procedures of Data Collection

The data was collected from 10 subjects (5 male and 5 female). Their age ranged from 17 to 22 years. The recordings were made in a silent zone where there was no background noise. ESL learners were given 50 indexing cards written on them the token phrase in between the token words, all ten vowels were placed as illustrating Table 1. Ten vowels were selected as follows: /iː, ɑː, ɪ, ɔː, æ, ə, uː, ɒ, e, ʊ/. Then five words from each vowel sound and then written them on the cards. Then we have given those cards to our respected subjects and then well requested them to pronounce each word three times. Each subject pronounces the 50 words three times like I say token word.

2.2 Speech Material

Five words from each vowel sound were chosen for the experiment. Total chosen words were mono-syllable words. The words frequency in daily use was highly familiar to ESL learners. Table 1 illustrates the fifty mono-syllable words which were used as the data tool as follows:

Table 1. Token words produced by ESL learners

Vowels	Words	Words	Words	Words	Words
iː	Peep	Sheep	Meet	Seat	Neat
e	Pet	Set	Get	Wet	Net
ɑː	Car	Far	Jar	Part	Half
ɔː	Core	Four	Snore	Score	Saw
ɒ	Pot	Cot	Short	Lot	Sort
ɪ	Fit	Bit	Sit	Pit	Kit
æ	Fat	Back	Bad	Cat	Sat
uː	Boot	Food	Soon	Moon	Tooth
ə	Dug	But	Bus	Cut	Bud
ʊ	Should	Could	Good	Put	Food

2.3 Recordings

Recordings were made of /hVd/ utterances recorded by ten male and female speakers. Measurements were analyzed of vowel duration, F0 contours (fundamental frequency), and formant frequency contours (F1-F2). A set of 50 cards with each word written on them was provided among the participating subjects. Total fifty words (25 for short & 25 for long vowels) were involved. The word structure was selected where each vowel sound was amid the word. There was a total of 50 cards, i.e., one card for each word inscribing "I say _token word_ now" in a phrase. In addition, the set of 50 cards was read with 03 repetitions by each ESL learners. The cards were reshuffled after each recording, and the subjects read aloud the word on each card in sequence with no break between repetitions. In addition, the participants were provided a quiet and calm environment, with no background noise, in a closed room far from the noise of city traffic. The voice recorder used faced them at a distance of around six inches. They were instructed to maintain a normal speech rate and to read the given words using a neutral focus, as in ordinary talking conditions. Subjects voiced were recorded on sound recorder and analyzed on computer Dell Core 2 dual system using the Praat speech processing tool (Boersma & Weenink, 2017) and a voice recorder.

2.4 Data Analysis

The recordings of voice samples were analyzed on Praat software for F1, F2, F0 and duration as basic parameters for the analysis. All three repetitive frequencies were calculated, and the average frequencies were taken as means of acoustic realizations of F1, F2, F0 and duration. The same vocalic measurement frequencies were on put on excel sheet for development of vowel space plots for male and female separately. After collecting the data from the subjects, we have analyzed it very carefully. We have analyzed every single word very carefully. We have taken F1, F2 three times of each word and then took their average. We have analyzed the duration and F0 of every subject very carefully. After taking all the frequencies from Praat software we have calculated their average and then we took the net averages of F1, F2, F0 and duration of every five words belongs to a vowel sound. Then we put the net values of F1, F2 and F0 in the J plot software and then we have got the plot of vowels.

3. Discussion

3.1 Measurements of Frequency F1 and F2

All measurements of frequency formants, pitch and duration of the vowels were taken manually by optical examination of wideband spectrographic display on computer screen. Start and end points of the target vowels were measured in milliseconds on the spectrographic displays. Both starting and ending points were restrained using the speech processing tool Praat (Boersma & Weenink, 2017). Frequency F1, F2 and fundamental frequency F0 of male and female were taken manually recognized on formant tracks at the visually positioned mid-point of the target vowel. Whenever a mismatch between the tracks and the visually apparent formant band in the spectrogram was detected, the formants were checked by visual inspection of wideband spectrographic display on computer screen. The pitch contours were manually extracted using the Praat autocorrelation method. Measurements were taken at the visually located mid-point of each target vowel. Duration was measured in milliseconds from the end of the preceding vowel to the stating. These measurements are illustrated graphically as below.
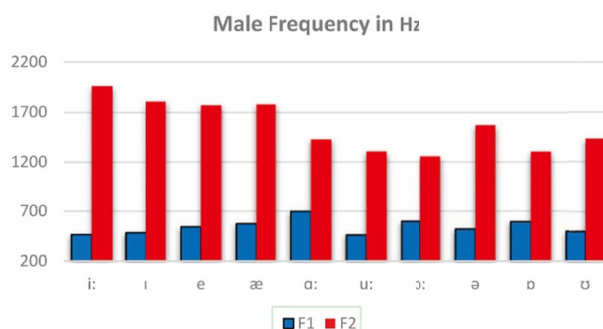

Figure 1. English ten vowels duration by across male ESL speakers

Figure 1 illustrates the frequencies F1 and F2 of the male subjects. Where the blue color represents the F1 and red F2. From this graphical representation which shows that the vowel ɑː has the highest value (698 Hz) of F1 and the vowel iː has the lowest value (464 Hz) of F1 among 10 vowels. The vowel iː has the highest value (1959 Hz) of F2 while ɔː vowel has the lowest value (1257 Hz) of F2 among the 10 vowel sounds.
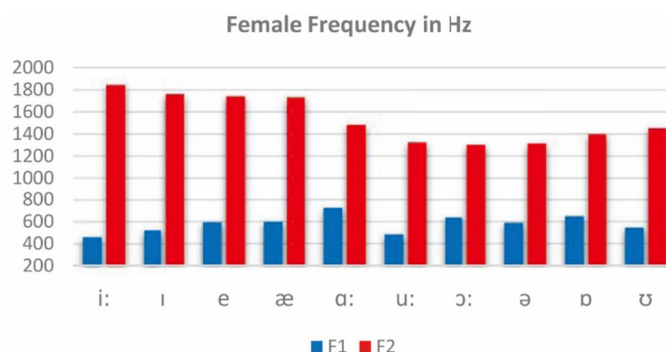

Figure 2. English ten vowels duration by across female ESL speakers

Figure 2 illustrates the frequencies F1 and F2 of the female subjects. Where the blue color represents the F1 and red F2. From this graphical representation which shows that the vowel ɑ: has the highest value (727 Hz) of F1 and the vowel i: has the lowest value (463 Hz) of F1 among the 10 vowel sounds. The vowel i: has the highest value (1846 Hz) of F2 while ɔ: vowel has the lowest value (1304 Hz) of F2 among the 10 vowel sounds.

3.2 Fundamental Frequency (F0)

All values have been taken from the same data. The data indicates that there is statistically difference in F0 values found between male and female speakers, however, there is a relatively significant difference between male and female speakers. The average F0 values were obtained from 10 speakers for all ten vowels. The recorded voice samples are represented in Figure 3 as follows:
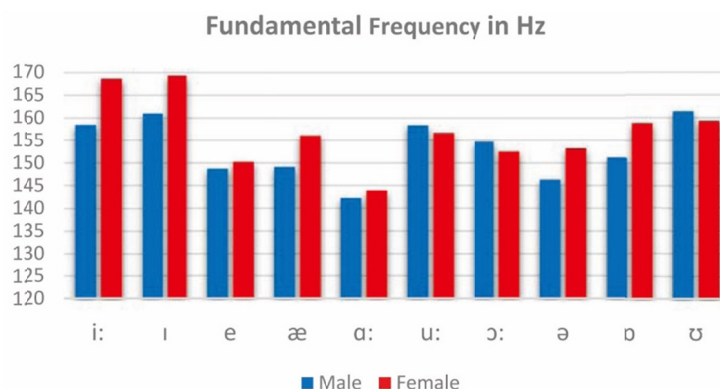

Figure 3. English ten vowels F0 by across female ESL speakers

Figure 3 illustrates the frequency F0 of the male and female subjects. Where the blue color represents the male and red female. From this graphical representation, the male subjects have the highest value (161 Hz) of F0 of ʊ vowel and they have lowest value (142 Hz) of F0 of **ɑ:** vowel among the 10 vowel sounds. The female subjects have the highest value (169 Hz) of F0 of ɪ vowel while they have lowest value (144 Hz) of F0 of vowel **ɑ:** amongst the 10 vowel sounds. Table 2 illustrates the paired sample statistics of male and female long vowel frequencies and their standard deviation and standard error including mean. Table 3 illustrates the paired difference statistics of male and female vowel frequencies and their standard deviation and standard error including mean and df 3.

3.3 Statistics of Long Vowels

Table 2. Paired sample statistics

Vowel Frequencies	Gender	Mean	N	Std. Deviation	Std. Error
F0	Male	155.425	4	10.2727	5.1364
F0	Female	153.525	4	7.6843	3.8422
Duration	Male	.4050	4	.01291	.00645
Duration	Female	.3875	4	.00500	.00250

Table 3. Paired difference

Vowel Frequencies	Gender	Mean	Std. Deviation	Std. Error Mean	95% Confidence	t	df	Sig. (2-tailed)
F0	Male	1.9	5.666	2.833	Lower -7.115	.671	3	.550
F0	Female				Upper 10.954			
Duration	Male	-.0175	.015	.0075	Lower -.0414	2.333	3	.102
Duration	Female				Upper .00637			

3.4 Statistics of Short Vowels

Table 4 illustrates the paired sample statistics of male and female long vowel frequencies and their standard deviation and standard error including mean. Table 5 illustrates the paired difference statistics of male and female vowel frequencies and their standard deviation and standard error including mean and df 5.

Table 4. Paired sample statistics

Vowel Frequencies	Gender	Mean	N	Std. Deviation	Std. Error
F0	Male	153.000	6	6.578	2.685
F0	Female	157.817	6	6.593	2.691
Duration	Male	.3850	6	.0164	.00671
Duration	Female	.3533	6	.0121	.00494

Table 5. Paired difference

Vowel Frequencies	Gender	Mean	Std. Deviation	Std. Error Mean	95% Confidence	t	df	Sig. (2-tailed)
F0	Male	4.817	4.152	1.695	Lower .4596	2.842	5	.036
F0	Female				Upper 9.1737			
Duration	Male	-.03167	.01722	.00703	Lower -.0497	-4.503	5	.006
Duration	Female				Upper -.0139			

3.5 Durational Values

The duration of ESL of vocalic sounds in some languages also noted by many phoneticians. The vocalic analysis of duration shows that there were statistically significant effects on vowel duration. The present study analyzed the duration of vowels and found long vowels were longer in duration, whereas, the duration of short vowels was shorter in duration relatively. The study also analyzed that duration of vowel sounds produced by female subjects were longer in duration and the vowel sound produced by male subjects were shorter in duration. Statistically overall mean shows that male speakers have the high value of duration as compared to female speakers. Male speakers produce the vowel sounds that take more duration as compared to the female speakers.
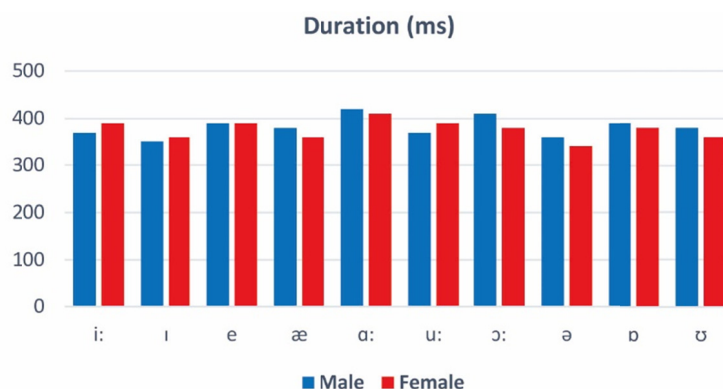

Figure 4. English ten vowels duration by across female ESL speakers

Figure 4 illustrates the time duration of the male and female subjects. Where the blue color represents the male subjects and red female subjects. From this graphical representation, it shows that the male subjects have the high duration (410 ms) of the vowel **e** and have the low duration (360 ms) of vowel **ɔ** among the 10 vowel sounds. The female subjects have high duration (410 ms) of the vowel ɑ: whereas the female subjects have low duration (340 ms) of **ɔ** vowel among the 10 vowels.

Ahmed (2005) discusses the production of vowels by Malaysian speakers that the short vowel has high average in terms of sound pitch and frequencies as compared to long vowel. They have found that the average pitch of the vowel sound is produced by female Malaysian speakers has high value as compared to male Malaysian speakers. Abbasi & Hussain (2015) found that there is significant difference between the stressed of short vowel of male and female and stressed long vowels of male and female and the unstressed of short vowel of male and female and unstressed long vowels of male and female. Neel (2008) has also found the result that adult has modified vowel space as compared to children. Abbasi, Chana, Kakepoto, Ali, & Mehmood (2017) argue that there is difference in terms of perception and production of English syllable and stress patterns while perceiving and producing English speech by Pakistani ESL learners.

4. Results

4.1 Overall F1 and F2 of Subjects

The result shown that the vowel sound produced by ESL speakers both male and female has different values of 4 research aspects (F1, F2, F0 and duration). The statistics which are shown in the graph shows that the average F1 of the male speaker has the highest value as compared to female speaker. The statistics shows that the average F2 of the female speakers has the highest value as compared to male speakers. The statistics illustrates that the average F0 of female subjects has the greater value as compared to male speakers. The statistics also shows that the male speakers have the high value of duration as compared to female speaker. The formant frequency (F1, F2) pitch and duration were analyzed by Praat the speech processing tool. From this graphical representation, which shows that the vowel ɑ: has the highest value (712.5 HZ) of F1 of males and female subjects and the vowel i: has the smallest value (463.6 HZ) of F1 of male and female subjects from 10 vowels. The vowel i: has the highest value (1902 HZ) of F2 of male and female subjects while ɔ: (1280.5 HZ) vowel has the smallest value of F2 of male and female subjects from 10 vowels.
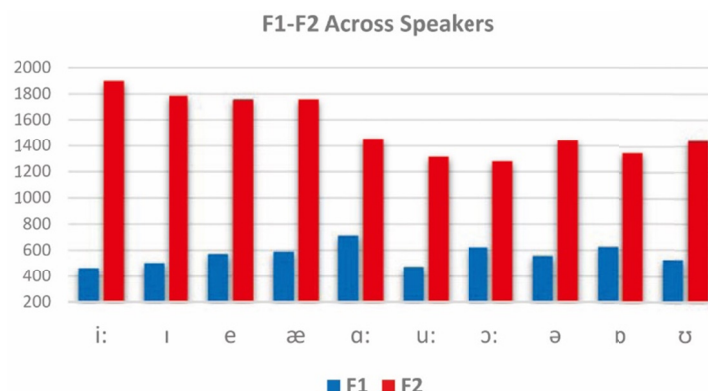

Figure 5. English ten vowels duration by across female ESL speakers

English vowel space plot illustrates ten vowels location within the vowel space produced by ESL learners as illustrated in Figure 6. The acoustic vowel space is based on first two formants; where horizontal axis corresponds to the gap between F1 formant and F2 formant whereas the vertical axis corresponds to the frequency of F1. Basically, it involves tongue body position representing high vs low and front vs back locations within the mouth cavity.
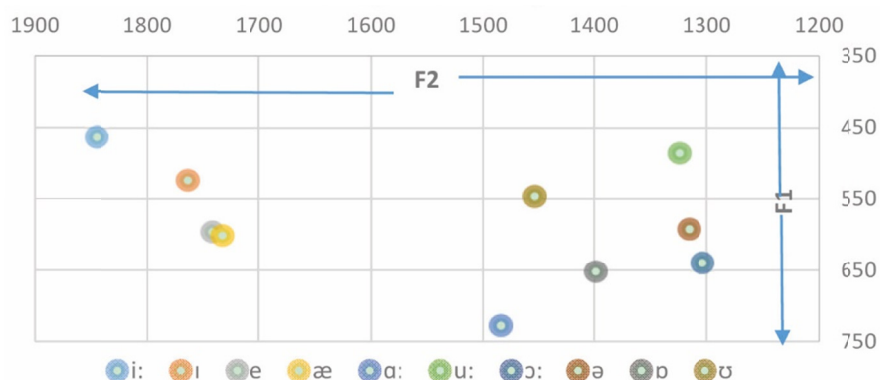

Figure 6. Female English vowel space plot

Female production of English vowels are as follows: Front vowel /i: / is located between 1800-1900 Hz F2 axis, and in height 450 Hz; whereas other front vowels i.e., /ɪ, e, æ/ are positioned between 1700-1800 Hz on the vowel space with the height of F1 500-600 Hz in between in mouth cavity. Central and back vowels /u:, ɒ, ɔ: and ʊ/ are located between F2 1200-1450 Hz horizontally and F1 while lower central vowel ɑ: decentralized towards the back.
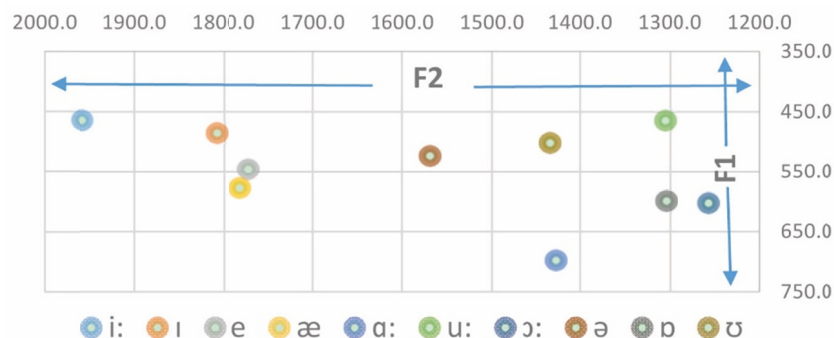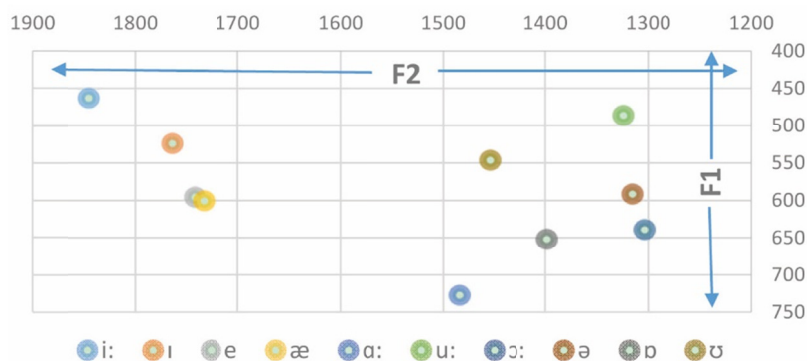
Figure 7. Male English vowel space plot

Figure 8. Overall English vowel space across speakers

5. Conclusion

The study concludes after the acoustic documentation of PaKE vowels properties that PaKE exists as a different variety spoken in Pakistan in terms of acoustic difference. All acoustic measurements were taken keeping in view of standards of acoustic-phonetic properties of speech sounds of any language in the world. Four parameters for the analysis of acoustic characteristics of English vowels were measured and found that F1 of the male subject has the higher value (713 Hz) as compared to female speakers between the groups. F2 of the female subject has higher value (1902 Hz) as compared to male speakers. F0 of the female speakers has higher value (170 Hz) as compared to male speakers. The duration of production of vowel sound of male subject is longer as compared to female subjects. Additionally, there are clear differences of English vowel space on vowel plots of PaKE produced by male and female ESL learners in terms of locating the production of vowels and the vowel space acoustic realizations.

References

Abbasi, A. M. (2010). *The production of English consonants by Sindhi ESL learners*. Munich. VDM Verlag Dr. Müller. Germany.

Abbasi, A. M. (2012). A phonetic-acoustic study of Sindhi-accented English for better English pronunciation. *International Journal of Social Sciences & Education, 2*(2)*, 146-157.

Abbasi, A. M. (2017). *The stress pattern of Sindhi and English*. Munich. *Lincom Europa*, Germany.

Abbasi, A. M., & Hussain, S. (2012). Syllable structure and syllabification in Sindhi-English loanwords. *International Journal of Researchers, 1*(4), 120-134.

Abbasi, A. M., & Hussain, S. (2015). Phonetic analysis of lexical stress in Sindhi. *Sindh University Research Journal, 47*(4), 749-756.

Abbasi, A. M., & Hussain, S. (2015). The role of pitch between stress and intonation in Sindhi. *Annual Research Journal of English Language Forum, 17*, 9-24.

Abbasi, A. M., & Kimball, A. (2014). Word stress in Sindhi and English: Implications for learners of English. In poster session presented at the meeting of *Sixth Annual Second Language Acquisition and Teacher Education (SLATE)* Graduate Research Symposium. Department of linguistics & foreign languages,

University of Illinois, Urbana-Champaign, USA.

Abbasi, A. M., Channa, M. A., John, S., Memon., M. A., Anwar, R. (2018). An acoustic investigation of Pakistani and American English vowels. *International Journal of English Linguistics, 8*(4), 115-121. https://doi.org/10.5539/ijel.v8n4p115

Abbasi, A. M., Channa, M. A., Kakepoto, I., Ali, R., & Mehmood, M. (2017). A perceptual study of phonological variations in Pakistani English. *International Journal of English Linguistics, 8*, 91-100. https://doi.org/10.5539/ijel.v8n2p92

Abbasi, A. M., Memon, M. A., Channa, M. A., & Stephen, J. (2018). Awareness of L2 American English word stress: Implications for teaching speakers of Indo-Aryan languages. *International Journal of English Linguistics, 8*(3). https://doi.org/10.5539/ijel.v8n3p101

Abercrombie, D. (1965). *Studies in phonetics and linguistics*. London: Oxford University Press.

Ahmed, W. A. S. (2005). Vowel length discrimination among Malay speakers of Malaysian English: An Instrumental Study.

Bauer, L. (2002). *An introduction to international varieties of English*. Edinburgh: Edinburgh University Press.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer (Version 5.3.35) [Computer software]. Retrieved from http://www.praat.org/

Crystal, D. (1999). *Dictionary of English and phonetics*. Cambridge: Cambridge University Press.

Dušan, N. (2016). Acoustic analysis of English vowels produced by American speakers and highly competent Serbian L2 speaker.

Ladefoged, P. (2001). *Vowels and consonants: an introduction to the sounds of languages*. Oxford: Blackwell.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *JASA, 35,* 1773-1781. https://doi.org/10.1121/1.2142410

McMahon, A. (2002). *An introduction to English phonology*. Edinburgh: Edinburgh University Press.

Neel, A. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*. https://doi.org/10.1044/1092-4388(2008/041)

Olive, J. P., Greenwood, A., & Coleman, J. S. (1993). *Acoustics of American English speech: a dynamic approach.* New York: Springer-Verlag. Review by T. Rossing in *Physics Today* March 1994, p. 59. http://dx.doi.org/10.1063/1.2808444/

Parsons, T. W. (1987). *Voice and speech processing.* New York: McGraw-Hill.

Rehman, M., Abbasi, A. M., & Lashari, A. A. (2017). Speech variations in Pakistani English and Singaporean English. *Journal of Social Sciences & Media Studies, 1*(1), 36-49.

Appendix A

Vowels	iː	e	ɑː	ɔː	ɒ
1	Peep	Pet	Car	Core	Pot
2	Sheep	Set	Far	Four	Cot
3	Meet	Get	Jar	Snore	Short
4	Seat	Wet	Part	Score	Lot
5	Neat	Net	Half	Saw	Sort
Vowels	ɪ	æ	uː	ə	ʊ
6	Fit	Fat	Boot	Dug	Should
7	Bit	Back	Food	But	Could
8	Sit	Bad	Soon	Bus	Good
9	Pit	Cat	Moon	Cut	Put
10	Kit	Sat	Tooth	Bud	Food

Copyrights