Journal of Social Sciences and Media Studies
Vol. 01 Issue No. 01

SPEECH VARIATIONS IN PAKISTANI ENGLISH AND SINGAPOREAN ENGLISH

Muneeb-u-Rehman[1] , Abdul Malik Abbasi[2]*, Ajab Ali Lashari[3]

Abstract

This paper presents acoustic analysis of six English vowels /a: /, /ɔ: /, /i: /, /ʊ/, /ɒ/, /u: / in three minimal pair of English words produced by second language learners in Pakistan. The paper aims to analyze acoustic realizations of Pakistani English vowels and compare them with Singaporean English vowels to investigate as to how much acoustic difference is in terms of F1 and F2 measurements between Pakistani English and Singaporean English. This study used the parameters of vowel quality (F1 and F2) Fundamental frequency 1 and 2 of English vowels. In addition, the analysis compares English vowels between the groups and within the groups. The data were collected through recordings of voice samples from nine subjects (5 male and 4 female) students. The study hypothesized that there are two varieties and will be different from each other, in terms of the production of Pakistani English vowels and Singaporean. The paper also looked at the male and female speakers' acoustic variations within the group recruited in Pakistan. The speakers were the undergraduate students from the department of Computer Science, Sindh Madressatul Islam University, Karachi. Speech differences between female and male voices are associated with multidisciplinary variations. Not only do these refer to acoustic (fundamental frequency) and perceptual judgments but also to the size of the vocal track in individual capacity coupled with educational background of the speakers. The study focused on the height and quality of the vowels to determine the acoustic differences between male and female students and Pakistani vs. Singaporean English.

Keywords: Acoustics, Speech Variation, Singaporean English, Pakistani English.

Introduction

English is spoken as an international language all over the world. It is spoken as the first language in more than fifty

[1]Mr. Muneeb Rehman, Student in Department of Computer Science, SMI University Karachi.
[2]*Dr Abdul Malik Abbasi, Corresponding Author, Assistant Professor of English, Allied Subjects, SMI University, Karachi. Email: amabbasi@smiu.edu.pk
[3]Mr. Ajab Ali Lashari, Lecturer Education, Department of Education SMI University, Karachi.

independent States or countries of the world. Almost half of the population of the world speaks English language as a source of communication. These days, English language plays vital role in term of science, medicine, arts, literature, business, engineering and most importantly computer and internet sciences. It is widely recognized as a language of trade and the language of science and technology, which meets the optimum level of reliability as the lingua franca (a language which is shared by non-native speakers). English is spoken in hundred different varieties and accents across the globe, many people speak different languages, in terms of English most people speak in different style like BBC (British Broad Cast) and RP (Received Pronunciation) etc. This study basically focuses on phonetic-acoustic differences between male and female frequencies objectively. The sounds of vowels are produced without obstacle, restriction and constriction of air mechanism from lungs through vocal track. These differences may be considered as a variety of pitches; however, they are also distinguished from each other by two characteristic pitches associated with their overtones (Ladefoged, 1993). One of them corresponds roughly to the difference between front and back vowels and the other corresponding to what we call vowel height in articulatory terms. These characteristic overtones are called the formants of the vowels. This study analyzed the acoustic parameters of vowels whose F1 frequency corresponds to the height of tongue and F2 corresponding to the back-ness of English vowels between male and female in terms of height (F1) and quality (F2) by Praat Speech Processing Tool.

Objectives of the Study

- Determine the acoustic differences between Pakistani English and Singaporean English.
- Investigate speech variations in terms of production of English vowels corresponding to (F1) and (F2) frequencies between male and female.

Statement Problem

Pakistani English is spoken differently accented English from Singaporean English. The problem statement of this work is to determine whether female spoken English is different from male English i.e., the quality (F1 and F2) of English vowels and to investigate the differences between these two varieties spoken in Pakistan and in Singapore.

Research Questions

i. What are the Acoustic variations in vowel quality of Pakistani and Singaporean (F1 and F2)?

ii. What are the Acoustic variations in vowel quality of Pakistani male and female (F1 and F2)?

Hypothesis

If the Pakistani learners of English produce English vowels, then there will be English speech variations in F1 and F2 height and quality of the speech in comparison to Singaporean English.

Literature Review

Languages differ in terms of not only its acoustic-phonetic corresponding but also its perceptual phenomenon. Each language contains its peculiarities regarding sound system of language. More work is available on acoustic analysis of native English vocalic sounds and the vowels of other languages. In so far as second language learners' production is concerned, there is not much work seen particularly in South Asian regions. Several researchers have really done work and their analyses in Pakistani spoken languages i.e., (Hussain, 2010 & Abbasi, 2015). However, some more scientific approach is required to look at the basic acoustic cues for the speech analysis. There are basically many parameters which correspond directly to the acoustic measurements i.e., Fundamental frequency (F0) or the pitch contours and pitch tracks, vowel quality (F1-F2). Researchers have also studied acoustic differences in male-female speech as a second language learners. F0 a fundamental frequency which is associated with the perceptual notion of pitch, is commonly regarded as the major difference between adult male and female voices. Takefuta et al. (1972) states that the mean F0 would be around 120 Hz for men and 200 Hz for women, however, these values slightly vary with the passage of age as argued by Pegoraro-Crook (1988) and become broadly lower for smokers as noted by Gilbert & Weismer (1974).

Previous studies have shown that there are cross-gender acoustic differences i.e., the vowel formants of female speakers tend to be located at higher frequencies (Hillenbrand, J., Getty, L.A., Clark, M., & Wheele, K, 1995; Pépiot, 2009). Takefuta et al., 1972 and Olsen, 1981, suggested that F0 range values are larger for female than for male speakers, even though there is no consensus on this point (see Simpson, 2009). Phonation type also seems to depend on the speaker's gender: female voices are often considered more breathy than male voices (Klatt & Klatt, 1990).

Fant (1966) states that the cross-gender acoustic variations can generally be accounted for by anatomical and physiological differences that arise during sexual maturity. Vocal cards become longer and thicker in male speakers (Kahane, 1978), that simplifies why they tend to vibrate more slowly than women. A second important anatomical issue is vocal tract length, that is, the distance from the vocal folds to the lips all things being equal, the longer the

vocal tract, the lower resonant frequencies (Fant, 1970). The normal length of the adult female vocal tract is about 14.5 cm, while the normal male vocal tract is 17 to 18 cm long (Simpson, 2009).

Acoustic-phonetic analysis of speech has been made practical by the advent of the speech spectrograph (Koenig, Dunn & Lacy, 1946), which has also encouraged a number of foundational questions regarding the perception of speech because spectrograms shown that speech is highly variable both within and between talkers. Among early researchers, Liberman, A.M., Cooper, F.S., Shankweiler, D and Studdert-Kennedy, M. (1967) focused on within speaker variation in the acoustic signs for stop place of pronunciation, while others focused on between speaker variations in the acoustic signs for vowels.

Helmholtz (1885) argues that speaker normalization refers to this research addressing on the fact that phonologically identical sounds show a great deal of acoustic variation across speakers, and that listeners are able to recognize words spoken by different speakers despite this variation. In defining speaker normalization in this way, we assume that phonological identity occurs when sounds are identified by listeners as orders of the same linguistic object (word or phoneme). For example, the word "cat" spoken by a man and a woman might be recognized. The importance of vowels formants (resonant frequencies of the vocal tract) in showing vowel sounds has been known for over a century, Helmholtz (1885) produced vowel sounds with resonators having frequencies that matched the vowel formant frequencies. The role of vowel formants in vowel observation was also confirmed by Fry, D.B., Abramson, A.S., Eimas, P.D. & Liberman, A.M. (1962) using a variety of artificial vowels.

In addition, Miller (1953) doubled the fundamental frequency of vocal fold vibration (F0) of two-formant vowels (from 120Hz to 240Hz) and found vowel category boundary swings for most of the vowels of English. Fujisaki and Kawashima (1968) also studied the role of F0 in vowel perception and found F1 boundary shifts of 100Hz to 200 Hz for F0 shifts of 200 Hz. Slawson (1968) predicted that an octave change in F0 produced a perceived change in F1 and F2 of about 10-12%. Listeners are also strongly affected by incompatible F0. Lehiste & Meltzer (1973) found lower vowel perception accuracy when they put children's high F0 with male vowel formants, and (to a lesser extent) when they put low male F0 with children's vowel formants. Gottfried & Chew (1986) found that listener vowel identification performance was less accurate when vowels were produced by a counter tone at a much higher F0 than is typical for a male voice.

It is well established that acoustic differences between male and female vowel systems are the result of both biophysical and socio-phonetic factors. However, there remains a good deal of uncertainty and storm about which

differences are to be attributed to which set of factors, and there are areas of disagreement within the different biophysical reasons which have been offered. (a) The female vowel system has a larger acoustic area. (b) There are non-uniform acoustic differences between male and female vowels. In specific the acoustic differences between male and female tokens of the same vowel category increases with formant magnitude: the F1 difference rises with vowel openness while the F2 difference rises with vowel frontless. By contrast, differences between the male and female back vowels /o: / and /u: / are slight. Average differences in male and female vocal tract geometry have gone some way to accounting for non-uniformity but have failed to resolve the magnitude of the comparable differences in vowel space size which have been found. So, for instance, Fant suggested that female speakers lower the formant values of the back vowels [o:] and [u:] thus bringing them closer to male values by using tighter and longer dorso-velar and labial strictures, exploiting the double Helmholtz resonate or like properties of these vowel categories. The socio phonetic aspect in vowel space magnitude differences becomes clear from cross-linguistic comparison. Although the larger female vowel space has been frequently found in data from different languages, the size of the difference is by no means constant.

Voice quality is a term that considers wide range of possible meanings (Abercrombie, 1967; Laver, 1980). Women and children have been somewhat mistreated groups in the history of speech analysis by machine. One reason is that most acoustic studies tend to focus on formant frequencies as signs to phonetic contrasts. The higher fundamental frequencies of women and children make it more difficult to estimate formant-frequency locations. Additionally, informal observations clue at the possibility that vowel varieties obtained from women's voices do not conform as well to an all-pole model, due perhaps to tracheal coupling and source/tract interactions (Fant, 1985; Klatt, 1986b). Examinations of the physics of larynx behavior (Stevens, 1981) suggest that the possible modes of sound generation fall into a small number of natural categories. Similarly, cross-language comparisons of phonemic contrasts comprising different laryngeal modes suggest the existence of only a few distinguishing contrasts (Ladefoged, 1973). According to Ladefoged, languages use the larynx in three typical ways: (1) by varying laryngeal tension so as to produce changes in fundamental frequency of voicing; (2) by adjusting the separation between the arytenoid cartilages to realize different phonation types such as glottal stop, creaky voice, modal voice, breathy voice, voiceless, and fully spread for breathing (see also Catford, 1964, 1977; Halle and Stevens, 1981; and Laver, 1980, for similar categorizations of phonation types); and (3) by varying the timing of the onset of voicing relative to supra-glottal articulatory movements to realize, for example, pre-voiced, voiceless-unaspirated, and voiceless-aspirated consonants.

Ladefoged proposes a set of multivalued distinctive features to capture linguistic contrasts along each of these continua and provides examples of languages that use each category distinctively. A similar set of laryngeal states has been identified by Halle and Stevens (1981) and described using binary distinctive features, called spread glottis, constricted glottis, stiff vocal cords, and slack vocal cords. The best set of distinctive features for characterizing the phonological/ physiological behavior of the larynx continues to be an area of some controversy. For our purposes, though, it is sufficient to note the contras five use versus normal vowels in languages such as Jalapa Maxatec (Kirk et al., 1984), the phonemic use of glottal-stop/glottalization gestures in Danish, or laryngealization as one of the characteristic properties of tone 3 in Mandarin Chinese, and the contrastive use of breathy versus normal vowels in languages such as Gujarati (Pandit, 1964; Fischer-Jorgensen, 1967). There is also the related use in some languages, such as Hindi, of voiced-aspirated stops, such as [bh], which are characterized by an interval of simultaneous voicing and aspiration following release (Dixit, 1989). To further in this research paper tries to find differences in male and female spoken English and to determine whether female speakers produce vowels are separate variety as compared to male varieties of English.

Methodology

Sampling

Nine participating subjects (five male and four female) were recruited from SMIU Karachi for collection of data recordings on laptop. They recorded their 162 voice samples on Speech Processing Tool Praat Software (64-bit edition: Praat 6.0.19 by Paul Boersma and David Weenink, 2016). The speakers were under graduate students of Computer Science Department Sindh Madresstul Islam University Karachi. Their age ranged from 19 to 25 years. Whereas, the Singaporean English speaker's recordings of the same vowel variation measurements was downloaded from internet (Low, E. L, 2015) in order to compare with vocalic acoustic realizations of vowel variations of Singaporean English and Pakistani English.

Speech Material

Speech material consists of a single carrier phase, which were six vowels tokens (/a: /, /ɔ: /, /i: /, /ʊ/, /ɒ/, /u:/) / hɑːt/, /hɔːl/, / hiːd/, / hʊk/, / hɒt/, / huːt/ used as voice samples in between the carrier phrase as vowel minimal pairs as follows:

I say *heart* now. I say *hall* now.

I say *heed* now. I say *hook* now.

I say *hot* now. I say *hoot* now.

Procedure

The subjects were given A4 size paper hand written for recording purpose. They were given a brief explanation as to how they could record their voice samples. These tokens were recorded on Praat Speech Processing tool on laptop. They were instructed to read aloud three times each carrier phrase. All speakers recorded their three repetitions on each voice sample of six tokens. They were also instructed to keep the microphone away three to six inches. The voice samples were recorded in noiseless room of computer laboratory SMIU Karachi.

Data Collection

The data were collected through recordings of second language learners' English speech of five male and four female students. The data of Singaporean English was collected through internet from NIE corpus link (Low, E. L, 2015).

Data Analysis

Data were analyzed with certain steps in a sequence. Firstly, the location of F1 and F2 was determined manually utilizing spectrographic red dots on Praat as point of F1-F2 values. The curser was kept on the first and second red dots on the spectrographic horizontal lines. However, the measurement was taken on the mid vowel keeping curser on it. The measurements were also checked on machine auto-check system navigating through options for their values. It was found that all speakers (male and female) had different F1 and F2 for English vowels. Thirdly, summary of each speaker (male and female) was taken from excel file so that their individual differences can also be seen and noticed. Finally, the data of four female speakers were measured statistically by applying *t-test* and compared with five male speakers. All English vowels were compared, and to find their probability *p-value*. The researcher took minimum three tokens of each sound in both varieties of English (male and female). Thus, total 162 voice samples (06 English vowels x 03 repetitions from each person x 09 speakers) for almost each English vowel was taken, measured, and compared with each other. The results are summarized as follows:

Table 1. Six English vowels mean values of F1 across male and female speakers

TABLE- 1 F1 values				
Vowels	Sound	Male Speakers	Female Speakers	P - Value
Heart	/a:/	672.121 Hz	843.855 Hz	
Hall	/ɔ:/	680.03 Hz	747.694 Hz	
				0.02
Heed	/i:/	388.745 Hz	424.047 Hz	
Hook	/ʊ/	472.800 Hz	532.362 Hz	
Hot	/ɒ/	684.849 Hz	746.845 Hz	
Hoot	/u:/	474.550 Hz	513.903 Hz	
Total Mean :		562.1826	634.7843	

F1 values are compared in between male and female speakers and found that female F1 values are higher than the male speakers. *P-value* was considered significance at $p < 0.05$. This study shows that the results are highly significant. The data is based on the voice sample of Pakistani English.

Figure 1. Bar chart illustrates the data of F1 across male and female speakers

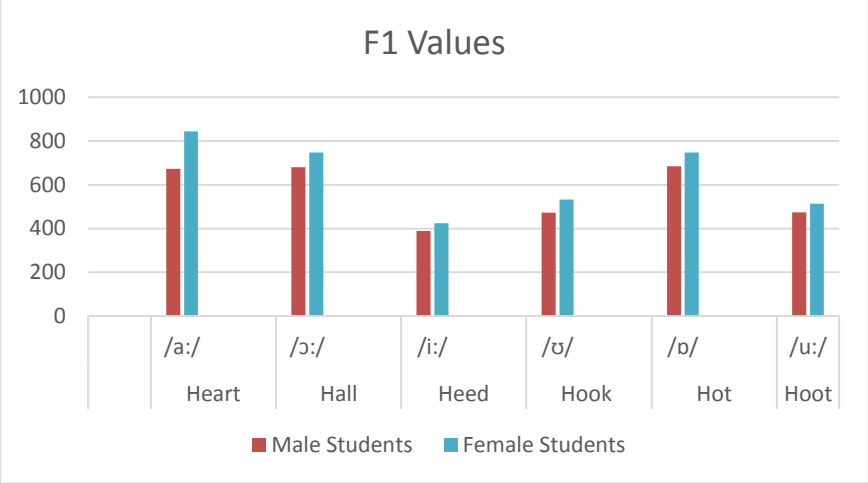
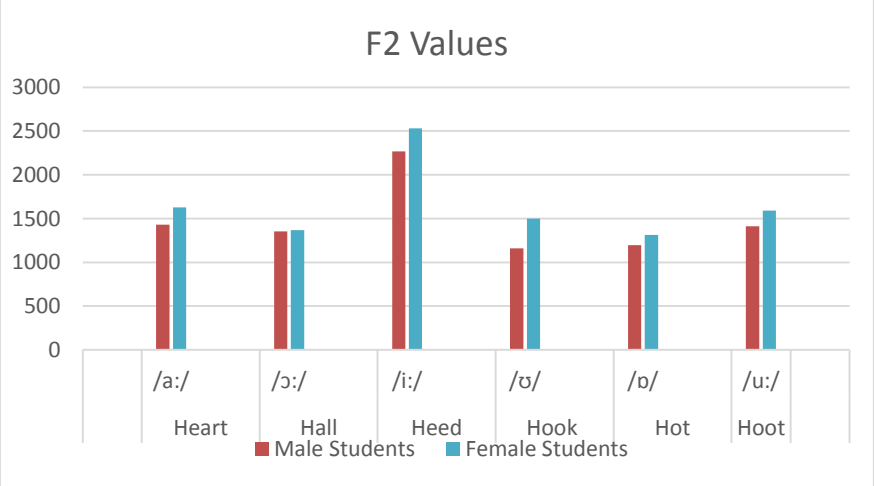


Table 2. Six English vowels mean values of F2 across male and female speakers

TABLE- 2 F2 values				
Vowels	Sound	Male Speakers	Female Speakers	P- value
Heart	/a:/	1429.306 Hz	1628.21 Hz	
Hall	/ɔ:/	1353.023 Hz	1367.942 Hz	
				0.01
Heed	/i:/	2266.491 Hz	2530.676 Hz	
Hook	/ʊ/	1157.952 Hz	1501.391 Hz	
Hot	/ɒ/	1196.755 Hz	1314.391 Hz	
Hoot	/u:/	1412.647 Hz	1590.424 Hz	
Total Mean		1469.362	1655.506	

F2 values are compared in between male and female speakers and found that female F2 values are higher than the male speakers. P-value was considered significance at $p < 0.05$. In this study, the results show that it is highly significant. The data is based on the voice sample of Pakistani English.

Figure 2. Bar chart illustrates the data of F2 across male and female speakers



There are more than hundred varieties of English spoken throughout the world. Languages differ in terms of speech variations. This may be phonetic, acoustic and phonological variations. This study has analyzed physical properties of English speech by Pakistani English and Singaporean English. The results of Pakistani and Singaporean speakers are summarized as follows:

Table 3. Six long vowels mean values of F1 across male and female speakers

Table 1 - F1 Values				
Vowels	Sound	Pakistani speakers	Singaporean speakers	P Values
Heart	/aː/	672 Hz	364 Hz	
Hall	/ɔː/	680 Hz	648 Hz	
				0.358
Heed	/iː/	389 Hz	386 Hz	
Hook	/ʊ/	473 Hz	449 Hz	
Hot	/ɒ/	685 Hz	796 Hz	
Hoot	/uː/	475 Hz	385 Hz	
Total Mean		562	505	

F1 values are compared in between Pakistani and Singaporean speakers and found that there is less significant difference. P-value was considered as significance at $p > 0.05$. In this study, the results show that it is not significant different.

Figure 3. Bar chart illustrates the data of F1 across Pakistani and Singaporean speakers of English

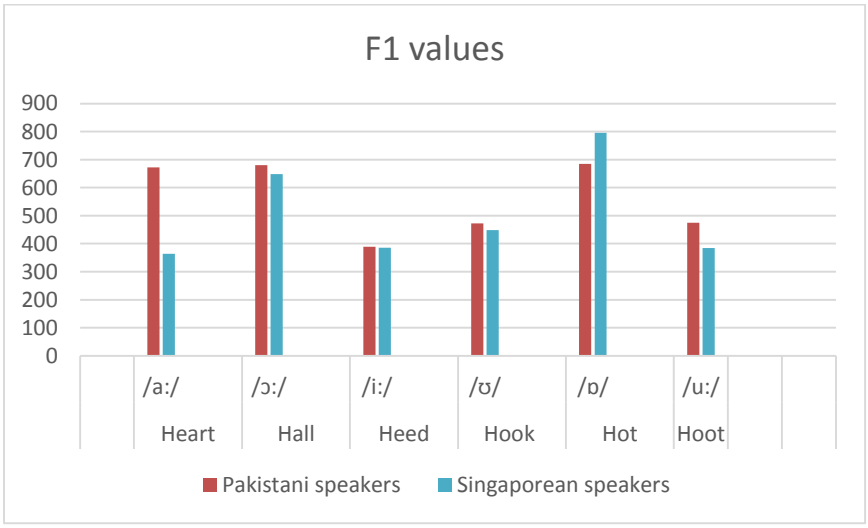
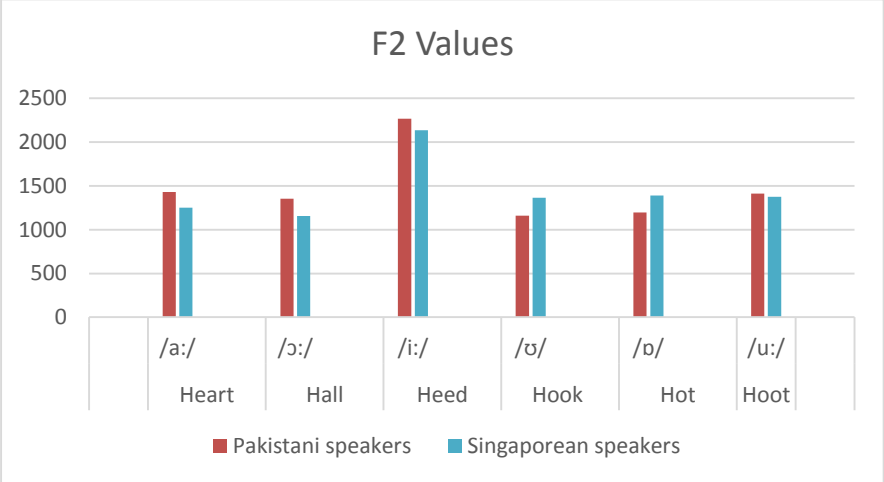


Table 4. Six long vowels mean values of F2 across male and female speakers

Table 2 - F2 Values				
Vowels	Sound	Pakistani speakers	Singaporean speakers	P Values
Heart	/a:/	1429 Hz	1250 Hz	
Hall	/ɔ:/	1353 Hz	1156 Hz	
				0.76
Heed	/i:/	2266 Hz	2134 Hz	
Hook	/ʊ/	1158 Hz	1366 Hz	
Hot	/ɒ/	1197 Hz	1389 Hz	
Hoot	/u:/	1413 Hz	1377 Hz	
Total Mean		1469	1445	

F2 values are compared in between Pakistani and Singaporean speakers and found that there is less significant difference. P-value was considered as significance at $p > 0.05$. In this study the results show that it is not significant.

Figure 4. Bar chart illustrates the data of F2 across Pakistani and Singaporean speakers of English



Discussion

As noted in the paper that the study analyzed the quality of female speakers' vowels and to highlight the difference between female and male speaker's English vowels in terms of their F1 and F2 frequencies. During the analysis, it was assumed that female and male students would be different from each other in their speech variations in Pakistani English. After the analysis, it was found that there was major difference in frequency 1 and frequency 2 between the

two (male and female speakers). It was also found that Pakistanis speakers as compared to Singaporean speakers do not make much different quality of sound from Singaporean speakers.

Conclusion

To conclude the data of Pakistani speakers and Singaporean speakers, it is claimed that Pakistani English vowels are not completely different from Singaporean English vowels. Though there are differences in the values of F1and F2 of these sounds, but the probability values show that there is no much difference in the two varieties of English language. In so far as male and female speakers are concerned, Pakistani English speech has highly significant difference in the production English vowels.

References

Abbasi, A. M, & Hussain, S. (2015a). Phonetic analysis of lexical stress in Indo-Aryan language. Sindh University Research Journal-SURJ, 47 (4), 749-756.

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh University Press.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.19) [Computer software]. Retrieved from http://www.praat.org/

Catford, J. C. (1977). *Fundamental problems in phonetics*: Edinburg: Edinburg University Press.

Cooper, W. & Sorensen, J. (1981). *Fundamental frequency in sentence production*, New York: Springer Verlag.

Dixit, R. P. (1989). Glottal gestures in Hindi plosives. *Journal of Phonetics*, 17, 213-237.

Fant, G. (1985). Speech technology - research and development. Presentation at the occasion of the Ericsson Prise 1985, *Ericsson Review*, 99-107.

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scaling. Speech Transmission Laboratory, *Quarterly Progress and Status Report*, 7, 22-30.

Fant, G. (1970). *Acoustic theory of speech production*, The Hague: Mouton.

Fischer-Jorgensen, E. (1967). Phonetic analysis of breathy (murmured) vowels in Gujarati, *Indian Linguistics* 28, 71-139.

Fry, D.B., Abramson, A.S., Eimas, P.D. & Liberman, A.M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech* 5, 171-189.

Fujisaki, H. & Kawashima, T. (1968). The roles of pitch and higher formants in the perception of vowels. IEEE Transactions on Audio and Electro Acoustics *AU-16, 73-77.*

Gilbert, H.R., and Weismer, G.G. (1974). "The effects of smoking on the speaking fundamental frequency of adult women," J. Psycholing. Res. 3, 225–231.

Goldstein, U.G. (1980). An articulatory model for the vocal tract of the growing children. *Thesis of Doctor of Science*, MIT, Cambridge, Mass

Gottfried, T. L., & Chew, S. L. (1986). Intelligibility of vowels sung by a countertenor. *Journal of the Acoustical Society of America* 79(1), 124-130.

Halle, M., & Steven. K. N. (1971). A note on laryngeal features. Quarterly Progress Report, MIT Research Lab of electronics, 101, 198-213.

Helmholtz, H.L.F. (1885). *On the Sensations of Tone* (Translated by A.J. Ellis) (2nd ed.). New York: Longmans.

Hillenbrand, J., Getty, L.A., Clark, M., & Wheele, K. (1995). Acoustic characteristics of American English vowels. *Journal of Acoustical Society of America*, 97 (5), 3099- 30111.

Hollien, H. (1974). On vocal registers, *Journal of Phonetics*, 2, 125-143.

Hussain, S. (2010). Phonetic correlates of lexical stress in Urdu. (Doctoral dissertation, Northwest University USA). National Language Authority, Islamabad, Pakistan.

Kahane, J. (1978). A morphological study of the human prepubertal and pubertal larynx. *American Journal of Anatomy*, 151, 11-20.

Kirk, P.L., Ladefoged, J. (1984). Using a spectrograph for measures of phonation types in a natural language, *UCLA Working Papers in Phonetics,* 59, 102-113.

Klatt, D.H. (1986). A spreading-activation theory of retrieval in sentence production. Psychol*ogical Review, 93(3),* 283-321.

Klatt, D., & Klatt, L. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-857.

Koenig, W., Dunn, H. K and Lacy L. Y. (1946). The sound spectrograph. *Journal of the Acoustical Society of American*, vol. 18, pp.21-32

Low, E. L. (2015). The NIE Spoken Corpus of English in Asia (NIESCEA). Singapore: National Institute of Education, Nanyang Technological University.

Ladefoged, P. (1973). The features of larynx, *Journal of Phonetics*, I, 73-83.

Ladefoged, P. & Antoñanzas-Barroso, N. (1985). Computer measures of breathy phonation. *UCLA Working Papers in Phonetics* 61, 79-86.

Ladefoged, P. (1993). A course in phonetics (3 ed.). Harcourt College Publishers, New York.

Laver, J. (1980*). The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Lehiste, I. & Meltzer, D. (1973) Vowel and speaker identification in natural and synthetic speech. *Language and Speech* 16, 356-364.

Liberman, A.M., Cooper, F.S., Shankweiler, D and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.

Miller, R.L. (1953). Auditory tests with synthetic vowels. *Journal of Acoustical Society of America,* 25, 114-121.

Pandit, P.B. (1964). Nasalization, aspiration and murmur in Gujrati, *Indian Linguistics, 17*, 165-172.

Pépiot, E. (2009). The making of French vocalic triangles: The case of a woman's voice versus a man's voice, *Master thesis*, University Paris.

Pegoraro-Krook, M. 1. (1988). "Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis," Folia Phoniatrica 40, 82–90.

Takefuta, Y., et al. (1972). A statistical analysis of melody curves in the intonation of American English, in Proceedings of the 7th International Congress of Phonetic Sciences, Montreal, 1035-1039.

Olsen, C. (1981). Sex differences in English intonation observed in female impersonation. *Toronto Papers of the Speech and Voice Society*, 2, 30-49.

Simpson, A. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3, 621-640.

Slawson, A. (1968). Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. *Journal of the Acoustical Society of America*, 43, 87-101.

Stevens, K. N. (1981). *Constraints imposed by the auditory system on the properties used to classify speech sounds: Data from phonology, acoustics, and psychoacoustics*. In T. Myers, Laver, J & J. Anderson (Eds.), The ognitive representation of speech (pp. 61–74). New York: North-Holland Publishing Company.