



Published by
**Centre for Languages and
Translational Studies**

Pakistan Journal of Languages and Translation Studies

ISSN (Print) 2410-1230
ISSN (Online) 2519-5042
Volume 10 Issue 2 2022
Pages 11-39

Open Access

An Acoustic Investigation into Diphthongs of Pakistani Variety of English

Publication Details

Arshad Mahmood
National University of Modern Languages
Email: armehmood@numl.edu.pk
margalla458@gmail.com

Paper Received:
October 05, 2022

Paper Accepted:
November 20, 2022

Paper Published:
December 28, 2022

Abstract

Pakistani English (henceforth, PakE), like all other *Inner Circle*, *Outer Circle* and *Expanding Circle* varieties (Kachru, 1992) of English, has its own hues and shades, which, naturally, render it a different variety of English. This study reports the diphthongal variation found in PakE with the help of spectrographic analysis of the target speech sounds. Each of the five English closing diphthongs /eɪ/, /aɪ/, /ɔɪ/, /əʊ/, /aʊ/ was embedded four times in a sentence and measured about 10% away from the opening and closing consonants (if any) in the syllable to minimize their (consonants') influence on the sound under investigation. The required data was garnered from 8 study participants (n=8), with even gender distribution, studying in the Department of English Language and Literature, National University of Modern Languages (henceforth, NUML), Islamabad, the capital city of Pakistan. The token words embedded in the carrier sentences were recorded at the FM Radio Studio of the university. The recordings were digitized at a 44.1-kHz sampling rate which is the set standard rate for audio files. The acoustic analysis was carried out with the help of Praat (6.1.16 version) with a focus on the overall duration of each diphthong, and the



Published by Licensee CeLTS. Copyright: © the author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

frequencies of F1 and F2. It was found out that PakE, like the other *Outer Circle* varieties of English, is a distinct variety of world Englishes which is developing/has developed its own phonological norms. The analysis of the data suggests that two of the closing English diphthongs (/eɪ/ and /əʊ/) lack the required glide, which is the hallmark of a diphthong. As far as the issue of gender is concerned, the analysis of the data suggests that it is a key factor in determining the vowel quality.

Keywords: World Englishes, Three Circles, PakE, diphthongs, glide, acoustic analysis, formant

1. Introduction

The notion of world Englishes is not a new concept anymore since various researchers have already proposed a number of models of world Englishes in order to document and record the spread of these new shades of Englishes up to this point (Stevens, 1980; Kachru, 1982, 1992; Moag, 1992; McArthur, 1987; Schneider, 2003, 2007). The focus of these models has been mainly on the acquisitional perspective as well as the developmental perspective (Low, 2010). Whether it is Moag's (1992) model of the 'life cycle of non-native Englishes' or Schneider's (2003, 2007) more recent 'dynamic model of postcolonial Englishes' or even Kachru's (1992) 'phases of recognition' of world Englishes, all of them recognize developmental cycles of new varieties of English (Kirkpatrick, 2007).

Out of all these models, Kachru's *Three Circles* seems to have attracted the attention of researchers most. The Three Circles proposed by Kachru are *the Inner Circle, the Outer Circle, and the Expanding Circle*. According to Kachru, *the Inner Circle* varieties (ENL speakers) are norm-providing and here English standards are therefore determined by these speakers (Kachru, 1992). These are mainly the United States, United Kingdom, Canada, Australia, and New Zealand. As far as *the Outer Circle* is concerned, this includes India, Malaysia, Philippines, Singapore, Sri Lanka, Nigeria, and so on and this is where the 'institutionalized non-native varieties (ESL) in the regions that have passed through extended periods of colonization' are used (Kachru, 1992). Kachru terms them norm-developing countries. Finally, *the Expanding Circle* includes China, Indonesia, Japan, Korea, Saudi Arabia, Russia and so on and these are the countries where performance varieties of the language are used essentially in EFL contexts (Kachru, 1992). Where the countries falling within the first two proposed circles are norm-providing and norm developing respectively, the countries falling within *the Expanding Circle* are norm-dependent and, thus, are largely reliant on the norms of English usage as prescribed by *the Inner Circle* varieties (Kachru, 1992).

The main focus of Kachru's *Three Circles* model was on the sociolinguistic profile of its speakers. To narrow it down further, the *Three Circles* proposed by him 'represent

the types of spread, patterns of acquisition and the functional allocation of English in diverse cultural contexts' (Kachru, 1992).

Looking back over the past, one can see that Kachru produced a number of works dealing with Indian English and Asian Englishes in the 1980s. These include *The other tongue: English across cultures* (1992), *The Indianization of English: The English language in India* (1983), and *The alchemy of English: The spread, functions, and models of non-native Englishes* (1986). Through these and other works, Kachru not only challenged the orthodoxies of English studies, he also accentuated the centrality of the pluralism of English worldwide. In other words, he dismissed the idea that non-native varieties of English were flawed and stigmatized approximations of standard or native English. Kachru believed that a language is heavily influenced and altered by its immediate ecology where there could be many competing languages around it. According to Kachru (2005), *the English language, as any other present or earlier transplanted language, is facing its ecological karma, and is woven into the nativized webs of language structure and its functional appropriateness* (pp. 255–256).

Kachruvian paradigm of world Englishes has been taken up by a number of researchers to study different Asian languages with regard to their phonological, morphological, lexical, as well as grammatical aspects. These studies were conducted on Pakistani English, Indian English(es), Malaysian English, Singapore English, Brunei English, Hong Kong English, and Philippine English (Bolton, 2012; Kitagawa, 2012; Yin, 2014; Abbasi et al, 2018). These studies and many others like them suggest that there are indeed shades of English that differ from standard British English as well as General American English. PakE is one of those varieties of English that fall within the *Outer Circle* of World Englishes.

The present research is an attempt to find out how PakE, being one of the norm-developing varieties of English that fall within the *Outer Circle*, is developing its own norms in terms of diphthongs. It aimed at exploring if diphthongs in PakE behave as double sounds, accompanied by a glide, or a single sound like a monophthong. Another aim of the study was to find out the F1 and F2 frequencies of the selected diphthongs, the duration of each of them and how they (frequencies and duration) vary with regard to gender. This was achieved with the help of 8 study participants who were asked to read out 5 sentences where each of them contained one of the target diphthongs four times.

The study was guided by the following research question and the subsidiary questions:

- 1- Does PakE make use of the five closing diphthongs which Standard British English and General American English do or is it (PakE) a distinct variety of English, with regard to its diphthongal variation, according to the Kachruvian *Three Circles* paradigm?

- 1.1- What are the F1 and F2 frequencies of PakE?

1.2-What are different durational values of PakE diphthongs?

1.3-What impact does gender have on duration and F1 and F2 frequencies?

The literature review provided below contains a great of deal of jargon pertaining to Phonetics, Phonology and Acoustics that is crucial to the understanding of single (monophthongs) and double vowel (diphthongs) sounds, their articulatory mechanism with reference to the resonance they carry because of the vocal cord vibration as well as different configurations of the vocal cavity, their formants, and their acoustic features. Most part of the literature review carries discussion on diphthongs, their nature, and different researchers' conflicting views about their nature as being a single sound or a double sound accompanied by a glide. This ends with a brief mention of different studies carried out to study various aspects of diphthongs in different parts of the world at different points of time.

2. Literature Review

A large number of different configurations of the human vocal apparatus are capable of producing 200 different vowel sounds (Ladefoged, 2001) with a large spectrum of acoustic characteristics. Technically speaking, all vowels achieve these characteristics by the degrees of raising, lowering, retracting and advancing the tongue (Sweet, 1890) in addition to other articulatory factors. Such a highly sophisticated vocal apparatus, operating like a chamber, is also suited to generate high degree of resonance called formants. Vowels are inherently more resonant than their counterparts: consonants. This is because of the periodic vibration of the vocal cords which must be on during the production of vowels. Due to such periodic vibration of the vocal cords during the production of vowel sounds, they are mainly voiced sounds. Vowels are generally categorized into monophthongs, diphthongs and triphthongs.

Since vowels can't be studied in terms of their Place of Articulation (PoA) and the Manner of Articulation (MoA) like consonants, they have been analyzed by different researchers with the help of their height, frontness/backness, duration and so on and the role of formants has been extremely important in this regard. Given the importance of vowel formants, their ever-fluctuating trajectories as well as their rate of change (roc), it would not be out of context to discuss them here briefly.

Formants

A formant is generally defined as a resonance of the vocal tract (Pickett, 1998). They are the visible peaks of acoustic power in a diagram of the output spectrum (Brosnahan & Malmberg, 1970). When it comes to vowel recognition, it is Formant One (F1) and Formant Two (F2) which play the most important role though F3 is also a generally useful indicator to the lip rounding. Different degrees of distribution of acoustic energy caused by a number of articulatory postures, which heavily influence vowel quality, differentiate vowel phonemes from one another. These differences can be

methodically studied by examining the distinguishable formant patterns and their respective trajectories. The chunk of a spectrogram given below shows the first three formants of a vowel sound with the help of red arrows. This can be very clearly seen that all formants have a very strong concentration of acoustic energy.

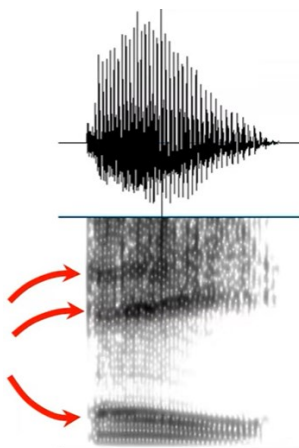


Figure 1: waveform and spectrogram showing formants

F1 and F2 correspond to height and frontness/backness of the tongue inside the vocal cavity respectively. In case the tongue body is held high (for example, for the vowel sound in ‘heat’), the back cavity is large, thus, rendering a low F1 value. On the contrary, when the tongue body is backed, the front cavity automatically becomes bigger and produces low F2. Keeping in mind the similar maneuvering of the tongue within the vocal cavity, we get high F2 when the front cavity is small due to fronting of the tongue body. In slightly easier terms, we will see an increase in the frequency with a vowel sound going from close to open. That is why, all high/close vowels have low F1 frequencies which keep shooting up with more open vowels. Likewise, an increase in the frequency of the second formant is easily observable with the vowel phoneme moving from back to front. For example, the English / I / is a front high close vowel and, naturally, its F2 frequency is higher than all the lower vowels in the system. The figure will explain the whole idea in a clearer manner. The red lines running over the dark bands point to the formant frequencies of the first five English monophthongs or single vowels. One can easily observe that the first horizontal band lies lowest because it belongs to the highest English vowel phoneme / I /. Comparing it with the last phoneme in the spectrogram, it is quite clear that the first formant has moved up because the said sound, which is /a/, is not as high as / I /.

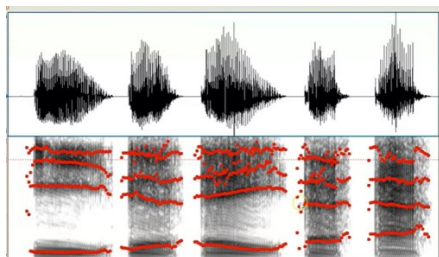


Figure 2: waveform and spectrogram showing formants of 5 English vowels.

Previously, it was believed that monophthongs have a relatively steady state as compared to their longer counterparts: diphthongs and triphthongs and it was claimed that this steady state played a very central role in measuring vowels since this is where researchers find rather precise measurements. Looking at this phenomenon in more physiological terms, we will realize that the steady state of vowels is achieved by stability and the fixed posture of speech organs during the very act of speech though the vocal folds are busy in their typical opening and closing during phonation or commonly known as voicing (repetitive opening and closing of the vocal ligaments occurs at different rates for males and females: approximately 125 Hz times per second in case of an average adult male, and almost 250 Hz in case of an average adult female. This faster cyclic pattern among females makes their voice sound high due to the high pitch) during the production of voiced sounds. Since the vibration of the vocal folds is one of the common hallmarks of vowel sounds, most of them (vowels) are voiced and we hardly find voiceless vowels in world languages.

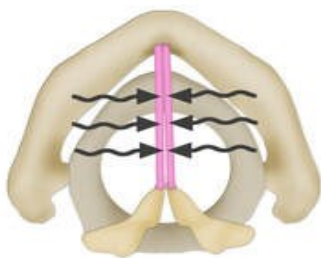


Figure 3: vocal cords

Source: <https://www.voicescienceworks.org/harmonics-vs->

Since most of the speech organs, except for the tongue and the vocal folds, are constant during the articulation of vowel sounds, such stability of the speech organs

naturally produces steady-state speech sounds which are very easy to be captured acoustically with the help of a spectrogram.

Nevertheless, the argument that a monophthong has a steady state has been challenged by many researchers who believe that no vowel sound is really static since spectral change is the inherent characteristic of each vowel sound (see, for example, Harrington and Cassidy, 1994; Nearey and Assmann, 1986).

Diphthongs

Tracing back the concept of diphthongs etymologically, we will find its origins in Greek 'phthong' which means 'vowel' or 'voiced sound'. This means a diphthong is a double vowel in which the articulatory mechanism moves continuously from an initial vowel position to a final vowel position (Kaiser, 1997). It is generally believed that all diphthongs are a transitionless glide between two short vocalic segments (McCully, 2009). According to McMahon (2002), they are essentially trajectories of articulation starting at one point and moving to another and this dual nature of them makes them appear like affricates. The vowel quadrilateral given below shows the gliding nature of a diphthong.

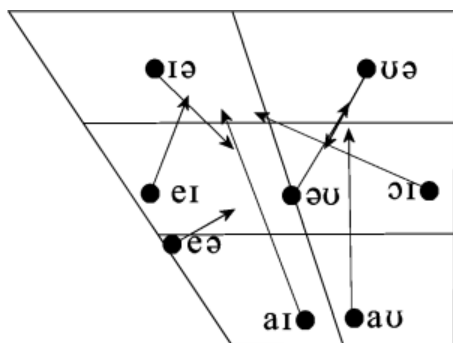


Figure 4: diphthongs and their gliding nature

However, the notion has complicated with the passage of time no matter how simple it appears in its original form. Considering the complex nature of these sounds due to their inherent dynamism, Lass (1984) rightly claims that 'Diphthongs are a positive migraine'.

Most probably because of such an intricate nature of these sounds, different researchers have understood the notion of 'diphthong' differently from time to time. Some take it as a single segment with an acoustically heterogeneous central chunk: segmental diphthong (Andersen, 1972), or it can also be a sequence of segments, falling within the same syllable: sequential diphthong (Andersen, 1972).

Debate whether a diphthong is a single event with an internal quality change or a phenomenon with two connected events has been going on for decades now. A number

of leading phoneticians (Lehiste and Peterson, 1961; Malmberg, 1963; Mose, 1964; Delattre, 1965; Abercrombie, 1967; Ladefoged, 1975; Bond, 1978; Jha, 1985; Lindau, 1990a; Gottfried, 1993) assert that it is one event with a continually changing quality. On the other hand, there are equally important phoneticians (Sweet, 1877; Jones, 1922; Hibbitt, 1948; Heffner, 1949; Romeo, 1968) who consider it a sequence of two vowels. There is yet another claim, though not supported and upheld by many phoneticians; the idea that a diphthong is a combination of one vowel and one semivowel (Trager and Smith, 1955).

The former school of thought believes in the duality (sequence) view (Arlund, 2007) and its proponents argue that diphthongs are two discrete simpler vowels phonetically connected to each other with the help of a smooth transition or the glide which is the result of an uninterrupted and seamless glide of the tongue within the vocal space. Conversely, the latter school of thought supports the unity (single vowel) view (Arlund, 2007), and the proponents of this school of thought contest that a diphthong, like any other phoneme of the system of a language, is a vowel sound and what distinguishes it from its counterparts (consonants and monophthongs) is its dynamic formant quality.

Even after such a hot debate, the argument does not seem to have come to a constant hold to date as there are phonologists who assert that a diphthong in one language may have features not found in diphthongs in other languages. This, naturally, makes it hard to reach a universal definition of this complex vocal notion (Lindau, 1990). Furthermore, there are experts who challenge this notion even in intralingual terms. For example, Pike (1947) suggests that [eⁱ, o^u, rⁱ, U^u] act phonetically as complex single phonemes (monophonemic) and [ai, au, ɔi] function as sequences of two phonemes (biphonemic).

Such acoustic variation in the case of diphthongs is most likely due to their dual role as a phonological event as well as a phonetic event in addition to some other important prosodic factors. What further aggravates the issue is the fact that similar diphthongs may behave differently cross-linguistically and this is surely caused by the marked differences among the prosodic variability as well as the multiplex nature of onglide, transition, and offglide found in these sounds of languages across the world. For example, an investigation by Lindau et al (1990) on American English, Peking Chinese, Egyptian Arabic and Hausa reports a lack of correlation between different transitional durations and the overall duration besides asynchronous movements of F1 and F2, for example in Arabic and Hausa. In the same vein, many language-specific studies have been carried out in the past (e.g., Aguilar, 1999; Bladon, 1985; Espy-Wilson, 1992; Jha, 1985; Wouters, 2002) to investigate the nature of these sounds. Holbrook and Fairbanks (1962), for example, report that the first part of a diphthong in English shows a bigger shift than the second part.

Another crucial issue in studying diphthongs is concerned with the point that what is the most important part of it to understand its nature best: onglide, glide or transition, or offglide, or, still, something else like the overall duration and/or the tempo. Gay (1968) seems to be the first person to study the effect of different speed rates on diphthongs. A similar investigation by Jha (1985) was carried on the Maithili diphthongs. Similarly, an important study on the acoustic and physiological parameters of diphthongs was conducted by Ren (1986). However, in terms of formant and their nature, the first study was brought about by Lehiste and Peterson (1961). Working further on, formants and their rate of change (roc) have been examined by some phoneticians (Kent and Moll, 1972; Manrique, 1979).

Besides the studies mentioned above, there have also been similar studies on different languages across the world. For example, a study by Kitagawa (2012) on the acoustic features of diphthongs produced by Japanese learners, studies on the dialectal nature of diphthongs in different parts of the world (Fox, & Jacewicz, 2009), acoustic analysis of vowels and diphthongs in Cairo Arabic (Norlin, 1984), dynamics in diphthong perception in American English (Jacewicz et al, 2003), vowels and diphthongs in Cangnan Southern Min Chinese Dialect (Hu & Ge, 2016), studies on the acoustic features of monophthongs and diphthongs (Elvin et al, 2016), comparison of diphthongs in Malaysian English and Singaporean English (Yin, 2014), an acoustic investigation of Pakistan and American English vowels (Abbasi et al, 2018) and so on.

This study aims to find out what closing diphthongs PakE as a norm developing variety of English uses. This will further strengthen the status of PakE as an *Outer Circle* variety of English.

3. Method

A. Speakers

All the study participants, aged 20-22 years, (4xmales;4xfemales=n:8) were enrolled in BS English Language and Literature program at the Department of English Language and Literature, NUML which is the only public sector university in the country with language teaching as its forte. None of the study participants spoke English as the Mother Tongue (MT) since English in Pakistan is taught and used as a foreign/second language. This study was undertaken with an aim to find out the acoustic features of the closing diphthongs (/eɪ/, /aɪ/, /ɔɪ/, /əʊ/, /aʊ/) to find out how PakE as an *Outer Circle* variety behaves diphthongally. Each diphthong was embedded in a carrier sentence four times in order to find out how it behaved each time in terms of F1, F2 and overall duration. At the end, there was a total of 160 (8x20) utterances to be analyzed acoustically and each one of them was analyzed thrice: for F1, F2 and duration (160x3=480). All the study participants were normal speech users since none of them reported any speech disorders. However, they were asked to rehearse the carrier sentences in order to overcome any potential fear. In case of a fumble or

hesitation, the recording was discarded and replaced by the correct one. All the study participants participated voluntarily and none of them was paid for the job; however, the teacher who supervised the whole activity thanked all the participants and offered them a cup of tea.

B. Recorded Material

Each carrier sentence contained 4 diphthongal words with a focus on the same diphthong. This is well-established now that the context (surrounding sound/s, lexical stress, sentence stress and intonation patterns) affects a sound in English. For example, the length of an English vowel is affected by the consonant sound which follows it and the same has been claimed by the stalwarts like Bloomfield (1983) who claims that English vowels are longer in the final position and before voiced sounds than before unvoiced or voiceless sounds.

The following carrier sentences were prepared for the sake of recording the required data:

- i. Dale's grey apron is on sale.
CVC-CV-VC-CVC
- ii. Giles's white tie is very nice.
CVC-CVC-CV-CVC
- iii. The hoity toity boy is enjoying his attitude.
CVC-CVC-CV-CV
- iv. The old man spoke like a croak due to a bad throat.
VC-CVC-CVC-CVC
- v. The stout boxer shouted in the bout in the last round.
CVC-CVC-CVC-CVC

Following table provides all the essential details of the token words.

Table 1. Token words

Diphthongs	Example words	Syllables	occurrence
ei	Dale, grey, apron, sale	CVC-CV-VC-CVC	4 times
ai	Giles, white, tie, nice	CVC-CVC-CV-CVC	4 times

ɔɪ	Hoity, toity, boy, enjoying	CVC-CVC-CV-CV	4 times
əʊ	Old, spoke, croak, throat	VC-CVC-CVC-CVC	4 times
aʊ	Stout, shouted, bout, round	CVC-CVC-CVC-CVC	4 times

Given that English sounds are acoustically affected by their surrounding sounds, the recorded samples were analyzed by truncating the onglide and offglide of each target sound almost 10% at the beginning and the end respectively. This means, the slice taken out of each diphthong for the analysis was the mid 80% which was almost unaffected by any phonetic or phonological context. The word list contained a variety of different CVC, CV and VC words. Each sentence, provided above, carried 4 words for the same diphthong.

C. Procedure

The said recordings were made in the NUML studio of FM Radio which has been set-up for airing different university programs and has quite a professional look. This is situated in the Iqbal Block of the university. A proper staff has been deployed by the university authorities to look after and run the whole setup. The staff provides teachers and researchers with all possible technical help. The researcher was able to get the data recorded in one session since there were only 8 study participants. In order to make them feel relaxed, they were offered a cup of tea and given ample time to rehearse the given sentences. The whole session went on smoothly and hassle-free data was obtained. The data was digitized at the standard 44.1-kHz sampling rate since it was an audio recording.

The spoken samples were drawn from real words of the language which were kept as free as possible from the determining effects of adjacent consonants. The collected data was analyzed with the help of spectrograms in order to measure F1, F2 and the overall duration of each diphthong.

All the closing vowels (/eɪ/, /aɪ/, /ɔɪ/, /əʊ/, /aʊ/) were embedded in 5 different sentences and measured about 20% away from the surrounding consonants (if any) in order to, as mentioned above, mitigate their influence on the sound under investigation because sounds are heavily influenced and affected by their phonetic context in English. In order to overcome this issue, some researchers have used the consonant/h/,

which does not exercise any significant influence upon the neighboring sounds, to flank their selected vowel/s. For example, Holbrook & Fairbanks (1962) used six surnames *Hay, High, Hoy, Hoe, Howe, Hugh* to address this issue. Similarly, Peterson & Barney (1952) employed *heed, hid, head, had, hod, hawed, hood, who'd, hud, and heard* though the ending sound did not behave like /h/ and its influence upon the preceding sound was managed otherwise. In this investigation, this problem was overcome with the help of truncation of each onglide and offglide by about 10% and it did serve the purpose.

4. Result

Eight study participants (4 male and 4 female) studying in the Department of English Language and Literature, NUML, Islamabad, were asked to read out 5 sentences that carried the token words with five closing diphthongs. The recordings were digitized at a 44.1-kHz sampling rate and the acoustic analysis was carried out with the help of Praat, 6.1.16 version with a focus on the overall duration of each vowel and F1 and F2 frequencies. All English closing diphthongs (/eɪ/, /aɪ/, /ɔɪ/, /əʊ/, /aʊ/) were selected and acoustically analyzed for the sake of finding out how PakE behaves in terms of these vowel sounds.

Analysis of collected data brings up some typical features of PakE which suggest that it is a variety of English that does not have a heavy reliance on diphthongs in their true sense: onglide, transition and offglide. It was observed that the articulation of a diphthong in some cases started off with an onglide but did not develop into a glide and an offglide, thus changing a bimoraic sound into a monomoraic or a pure sound', especially in the case of /eɪ/ and /əʊ/ sounds. With such a change to the pronunciation of a diphthong, there is naturally a strong impact on F1 and F2 frequencies and it is proven that the role of F2 is extremely vital in the investigation of diphthongs.

As far as duration is concerned, it was found that this variable was heavily affected by gender; female study participants produced these sounds with longer durations as compared with their male counterparts (see Appendix A & B). Likewise, there was also intragender variation but it is not a big issue because this is measured in milliseconds (ms) and a difference of some milliseconds is not significant at all. Variation in the length was shown in the form of either vowel shortening or vowel lengthening. Such shortening and lengthening could be the result of the Mother Tongue (MT) framework, sociophonetic factor and/or cross-sectional vocal tract dimension differences between males and females.

Given below are four spectrograms of the four words produced by a study participant. They have been shown in reduced sizes just to save space. They are followed by a collective spectrogram of the same diphthong, i.e., /aɪ/.

These four spectrograms show /aɪ/ diphthong whose first segment is the open, low sound and the second segment is the high, close sound. Since the distance between F1

and F2 displayed by / a / in English is fundamentally shorter than that of / ɪ / sound, the dark bands of the former can be seen starting off very close and then digressing from each other along the transition toward the latter sound where they diverge very obviously. This divergence is due to the fact that the English / ɪ / sound has the lowest F1 but the highest F2 (approximately 280 and 2250 for F1 and F2 respectively) among all monophthongs of the system (Ladefoged & Johnson, 2011). On the contrary, the F1 and F2 frequencies of / a / are approximately 830 and 1170 respectively. This means, the tongue glides from a relatively lower and more open point (because of the lower vowel / a /) to a higher and closer point (because of the lower vowel / ɪ /). This whole phenomenon is visible in the form of a formant trajectory which has a fork-like shape.

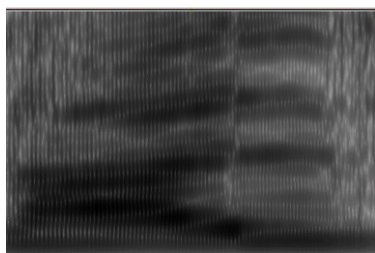


Figure 5: spectrogram of ‘Giles’

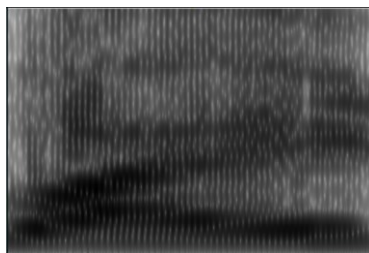


Figure 6: spectrogram of ‘White’

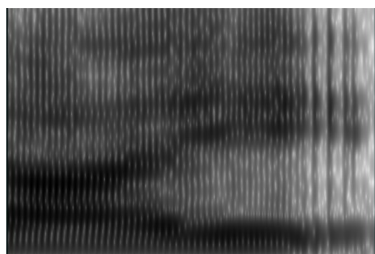


Figure 7: spectrogram of ‘Tie’

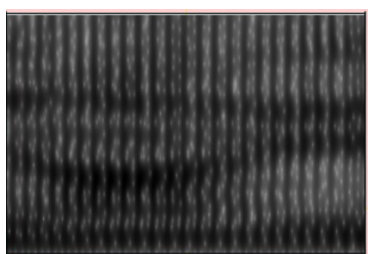


Figure 8: spectrogram of ‘Nice’

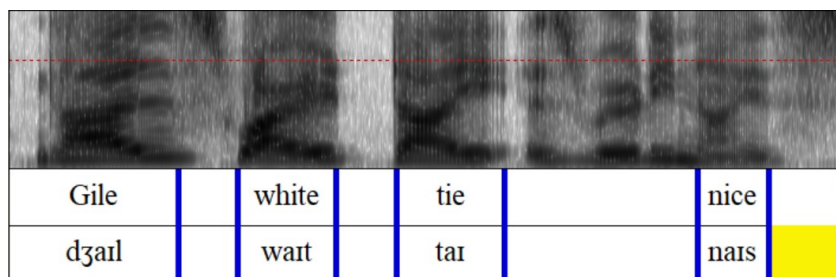


Figure 9: spectrogram of ‘Giles’s white tie is nice’

The spectrogram presented above shows how the sentence ‘Giles’s white tie is very nice’ was pronounced by one of the study participants. The top tier shows the

spectrogram of the said sentence, the text grid comprises two tiers where the first one shows the annotations of the given sentence in the normal English spelling and the second tier provides the phonemic transcription of the same. Other features of the spectrogram lying on the x-axis and y-axis have been removed on purpose.

The upward formant trajectory in the first three words, shown in the dark black bands due to the concentration of the acoustic energy can be traced very easily. The fork-like shape shows how F2 for each of the three words diverges quite far from F1 though their trajectorial patterns are not uniform. As far as the fourth word is concerned, its trajectory is not as clear as those of the first three words, yet one can easily notice the little divergence of F2 towards the offglide phase. These diverging formant patterns shown above depict the state of affairs very clearly. As evident, one can see that F1 and F2 start off their journey in very close proximity within the vocal tract due to the presence of the low vowel / a / and then diverge from this point due to the glide of the tongue caused by the presence of the high front vowel / i /.

The following graph shows the formant frequencies of all English monophthongs inside the vocal space. The graph has two dimensions where the horizontal dimension represents the frequency of the first formant (F1) and the vertical dimension represents the frequency of the second formant (F2).

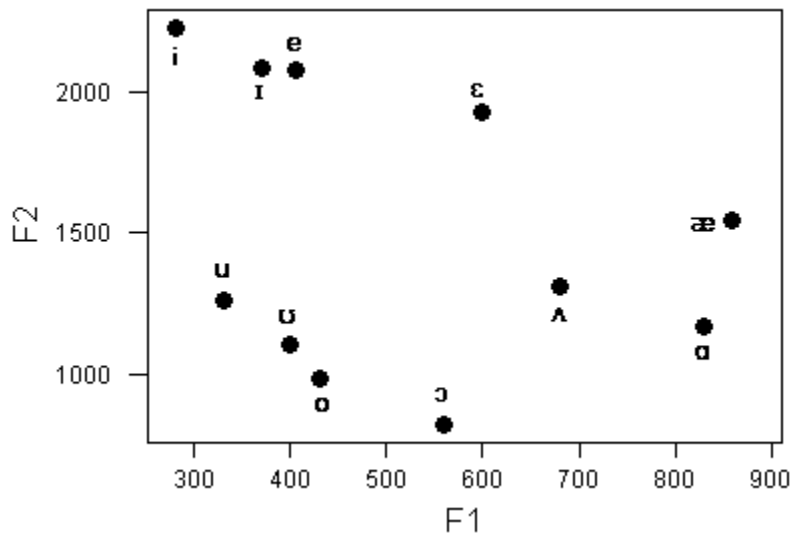


Figure 10: English vowels within vocal space and their formant frequencies

A pair of spectrograms is provided below to show how diphthong /əʊ/ was pronounced by one of the study participants in two different ways in two different words.

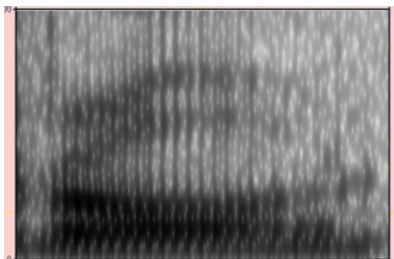


Figure 11: slice of ‘Spoke’

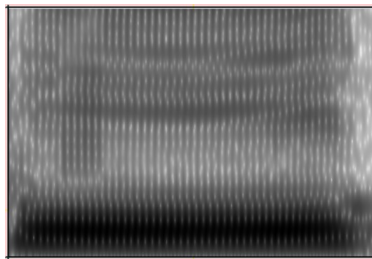


Figure 12: slice of ‘Throat’

The two spectrograms provided above (Fig. 11 & 12) show the dark bands caused by the production of ‘spoke’ and ‘throat’ respectively which share the same diphthong, /əʊ/. Here, the spectrograms paint two different pictures. In the first spectrogram, we see the slight divergence of the acoustic energy towards the end of the sound which is often seen in the spectrograms of /ɔ /sound in English but the onglide doesn’t seem to match that of a schwa sound (/ə /) where the space between the first two formants should be wider. The second spectrogram shows a very blurry picture where F1 and F2 seem to be cutting into each other. As a result of such conflation, it is hard to differentiate between the two formants in the second spectrogram. Since F1 and F2 for /ɔ / in English are approximately 560 Hz and 820 Hz respectively, the dark bands in the first figure lie very close to each other (though they lie too close to be the exact formants of /ɔ /). On the other hand, F1 and F2 for /ə / in English are approximately 665 Hz and 1772 Hz respectively but neither of the spectrograms shows either the onset for schwa sound (/ə /) or the required transition towards the second element (/ʊ /) due to which there seems to be no dynamism in the form of divergence which is the hallmark of a diphthong. There is another issue with F1 in both the figures; it lies very low in both of them whereas the first formant for /ə / as well as /ʊ / is/should be slightly higher when seen in a spectrogram. Monophthongization, or shortening, not only changes a double sound into a single sound, but it, in most cases, also reduces its duration. This vowel (/əʊ /) seems to have developed into a monophthong in PakeE which is, like the other *Outer Circle* varieties of English, developing its own norms.

Diphthong /əʊ /, in this study, which went through shortening reflected glideless movement from one point to the other. In other words, it did not carry any transition along and, therefore, sounded more like an /o / which is not a diphthong but a monomoraic sound in Urdu and is commonly found in Urdu words like *bolo* (say, speak) and *chor* (thief). But what is more interesting is the fact that the same diphthong appeared quite different on the spectrogram due to varying treatment by the study participants. That is to say, the same sound was produced as a diphthong, though very rarely, at one point but as a monophthong at another point. As a result of such monophthongization which most of the token words underwent, the transition was lost on the way and, as a result, no divergence was visible where F2 started off.

Given below is a spectrogram which shows another diphthong (/ eɪ /) pronounced four times by the same study participant differently.

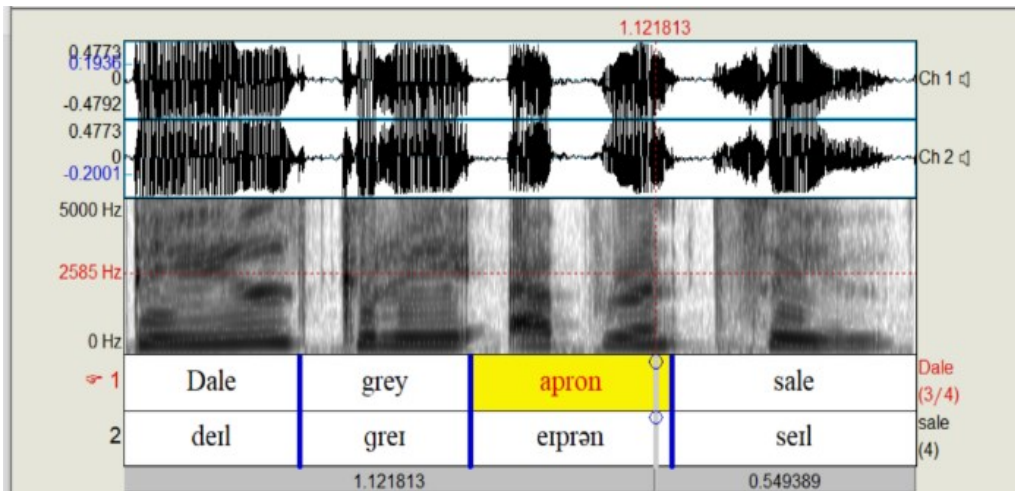


Figure 13: waveform, spectrogram and annotations

The spectrogram in the second tier in the grey shade shows the marked difference in both duration and glide in the case of the words ‘grey’, ‘apron’, and ‘sale’. However, the first word ‘Dale’ carries some diphthongal characteristics but this, in fact, is not what a spectrogram of / eɪ / looks like. If we look at a typical spectrogram of / eɪ /, we will see an onglide caused by / e /, a smooth glide towards / ɪ / and divergence of F1 and F2 caused by / ɪ / which has the lowest F1 (approximately 280 hertz for males and 310 for females) and the highest F2 (approximately 2250 hertz for males and 2790 for females) among all English monophthongs. The spectrogram given above shows quite a low F1 for the first two words as well as the last word, and a relatively high F1 for the ‘ap’ part of ‘apron’. As for the second formant, we see a slightly broken and blurry F2 in the first word though it is not as high as F2 of / ɪ / sound, an even blurrier F2 in the second word, a very low F2 in the third word and again a low F2 for the last word where none of them matches the height of F2 that we see in the spectrograms of / ɪ / (see Ladefoged, 2006; Peterson & Barney, 1952). To be precise, this English diphthong behaves more like the cardinal vowel [e] which can be found quite abundantly in Urdu words, for example, *sher* (lion) and *meray* (my). This suggests that / eɪ / largely changes into a monophthong by losing all its diphthongal characteristics due to one reason or another though it may put on some diphthongal behaviour at some places and contexts which need further research. The process of monophthongization which changes a diphthong into a monophthong also affects F1 and F2 frequencies of the sound besides altering its fundamental frequency (F0).

Besides shortening either in the form of monophthongization or otherwise, a tendency for lengthening in PakeE has also been observed. In this case, certain sounds are exaggeratedly pronounced and this blowing them up allows them to capture quite a noticeable place in acoustic terms in the spectrogram. An instance of lengthening is provided below where a spectrogram showing the words ‘boy’ and ‘enjoying’ has been taken out for a better understanding of how temporal space within the vocal tract is affected when a sound is elongated. This tendency was more commonly found among the female study participants.

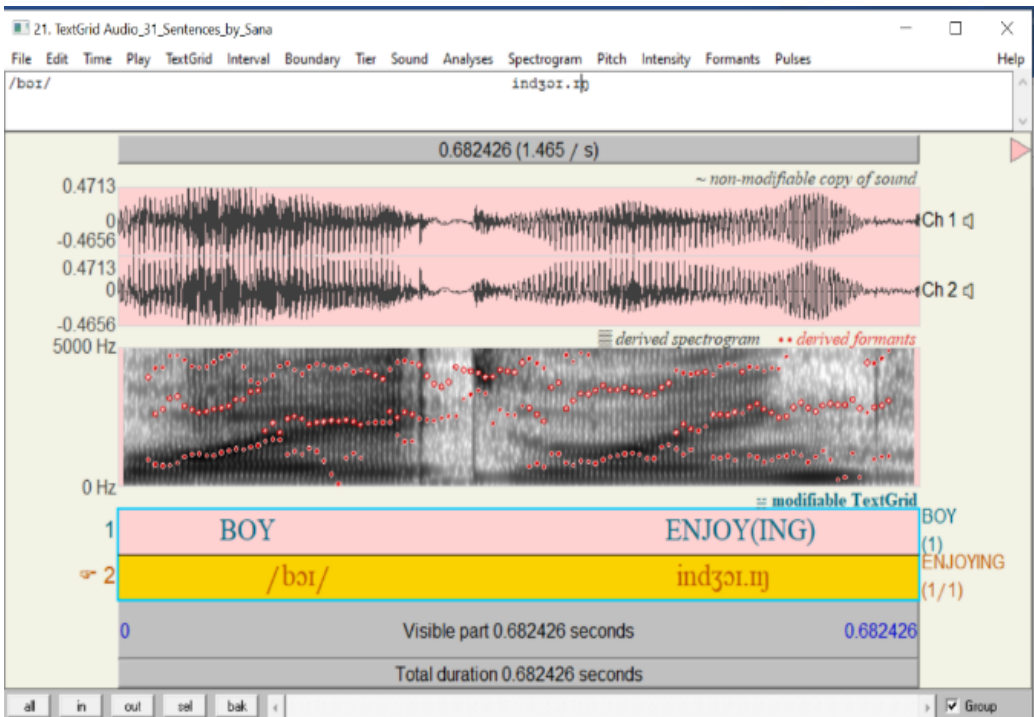


Figure 14: waveform and spectrogram and annotations

Diphthongs which were mostly monophthongized were mainly (/eɪ/ and /əʊ/ though their length might have been longer than their corresponding monophthongs of the system. The analysis of the data provided another interesting feature of PakeE. It was observed that though the diphthong /ɔɪ/ did not undergo monophthongization in most cases, its duration was very short in two of the given words: ‘hoity’ and ‘toity’.

As for the issue of gender, formant frequencies for vowels are different in different sex groups (Peterson & Barney, 1952) since the shape and size of the vocal apparatus of men and women are remarkably different. As discussed above, F1 and F2 frequencies correspond to the tongue height and the tongue backness/frontness respectively, there are potential anatomical and physiological differences among men and women when it

comes to the production of speech sounds. In fact, vowel formant frequencies are affected by a number of factors including the intrinsic size of the vocal tract, the size of the tongue, the size and configuration of the oral cavity, the size and configuration of the pharyngeal cavity, and the tongue configuration (Liu & Ng, 2009).

Considering the issue of gender in this study, the analysis of the data suggests that female participants produced relatively longer sounds as compared to their counterparts. This supports the findings of Simpson & Ericsson (2003) and Ericsson & Ericsson (2001). This means, they used more temporal space inside the vocal tract than the male participants did. Besides duration, the F1 and F2 frequencies for all the selected diphthongs were found to be higher among the female study participants as compared to their male counterparts (see Appendix A & B). This finding of this research supports the findings of Kwon (2010). These durational and frequency-related differences between the male and female study participants could be due to the Mother Tongue influence, anatomical and physiological differences in the vocal tract, the size of the vocal folds(which surely affects the vibration rate),the size of the tongue and the vocal apparatus, cross-sectional vocal tract dimension differences between males and females as well as sociophonetic differences as women have been reported to speak more clearly than their male counterparts (O. Elyan,1978).

The table given below provides the Mean Values for F1 and F2 and duration according to gender.

Tables 2: Mean Values of F1, F2 & duration -all study participants

Sound	Mean F1	Mean F2	Gender	Mean (Duration)	Difference in Mean Duration
/eɪ/	470.44	2213.1	Female	231.81	/ eɪ / 31.87
/eɪ/	398.25	1793.8	Male	199.94	
/aɪ/	770.56	1979.2	Female	256.13	/ aɪ / 24.38
/aɪ/	554.19	1469.6	Male	231.75	
/ɔɪ/	625.38	1689.9	Female	214.44	/ ɔɪ / 15.25
/ɔɪ/	437.13	1471.1	Male	199.19	
/əʊ/	481.06	1203.3	Female	227	/ əʊ / 27.88
/əʊ/	461.5	1178.4	Male	199.19	
/aʊ/	722.63	1583.9	Female	219.63	/ aʊ / 17.32
/aʊ/	628.44	1459.2	Male	202.31	

5. Conclusion

Pakistani English is a distinct variety of English which is the result of the impact caused by internal as well as external factors. Internal factors are, perhaps, stronger in terms of their phonological influence on PakE as it is an established fact that one's Mother Tongue (MT) system has a very strong influence on one's pronunciation of a foreign language. This is because of the fact that the learner/user of a foreign language mostly employs his Mother Tongue (MT) framework to seek guidance and he/she has a very strong tendency to make an appeal to his/her Mother Tongue (MT) framework whenever he/she feels such a need.

The analysis of the data shows that PakE does not make any remarkable use of diphthongs though we may quite often find an increased length of a sound. This is an instance of monophthongization where a double sound changes or is changed into a single sound. On the whole, it has been observed that / eɪ / and / əʊ / lack the three typical phases in PakE that every diphthong should go through. These three phases are the *onglide*, the point at which the vowel sound starts and reaches a steady state, a *transitional phase* due to the smooth glide of the tongue body from one place to another, and the *offglide* which, in fact, is the termination of the sound. Out of these three, the first two have been reported by many researchers to be of prime importance since they are not affected by the factors like the tempo of the speaker (Gay, 1968; Jha, 1985).

One import finding that came up was the durational variation found in the same sound pronounced 4 times in a sentence. This suggests the lack of diphthongal consistency found in the Pakistani variety of English. However, such a durational variation was not very obvious in the pronunciation of /aɪ/ and /aʊ/ diphthongs. This is most probably the result of the presence of these two sounds in most Pakistani languages which means that study participants did not have to struggle with the articulation of these sounds like the other typical Received Pronunciation (RP) diphthongs. As for the articulation of words carrying / eɪ / and / əʊ / sounds, they were mostly monophthongized since these appear to be among the rare diphthongs missing even in some varieties of English as well (Thewlett & Beck, 2013).

Returning to Kachruvian *Three Circles* model which includes *the Inner Circle*, *the Outer Circle*, and *the Expanding Circle*, where the *Inner Circle* varieties are norm-providing countries, *the Outer Circle* varieties are norm-developing countries and *the Expanding Circle* varieties are norm-dependent countries, PakE truly belongs to *the Outer Circle*. The analysis of the data suggests that PakE, like all *the Outer Circle* languages, is developing its own linguistic norms instead of strictly following the patterns of the languages which fall within *the Inner Circle* which are the norm-providing languages. Since the focus of this study was on the diphthongs of PakE with regard to their durational values and F1 and F2 frequencies, it was found that PakE is an *Outer Circle* variety of English which makes use of only three (/aɪ/, /ɔɪ/ and /aʊ/)

diphthongs out of five closing diphthongs which Standard British English and General American English do.

Acknowledgments

This research was carried out under the supervision of Dr. Ryan Skinnell (Department of English & Comparative Literature, San José State University) for my postdoctoral studies in the United States of America in 2022. I am extremely thankful to Dr. Skinnell for his valuable support and guidance in this regard. I am also thankful to Ms. Afsheen Kashifa, lecturer at NUML, and the study participants for providing the required data.

References

- 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).
- Abercrombie, D. (1972). LF Brosnahan and B. Malmberg, *Introduction to Phonetics* (pp. ix+ 243. W. Heffer & Sons Ltd., Cambridge, 1970.). *Journal of the International Phonetic Association*, 2(2), 85-89.
- Aguilar, L. (1999). Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech communication*, 28(1), 57-74.
- Andersen, H. (1972). Diphthongization. *Language*, 11-50.
- Arlund, P. (2007). An acoustic, historical, and developmental analysis of Sarikol Tajik diphthongs.
- Bladon, A. (1985). Diphthongs: A case study of dynamic auditory processing. *Speech Communication*, 4(1-3), 145-154.
- Bloomfield, L. (1983). An introduction to the study of language. *An Introduction to the Study of Language*, 1-383.
- Delattre, P. (1964). Comparing the vocalic features of English, German, Spanish and French.
- Elvin, J., Williams, D., & Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the acoustical society of America*, 140(1), 576-581.

- Ericsson, C., & Ericsson, A. M. (2001). Gender differences in vowel duration in read Swedish: Preliminary results. Working papers/Lund University, Department of Linguistics and Phonetics, 49, 34-37.
- Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels/wjrl/in American English. *The Journal of the Acoustical Society of America*, 92(2), 736-757.
- Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America*, 126(5), 2603-2618.
- Gay, T. J. (1967). *A Perceptual Study of American English Diphthongs* (Doctoral Dissertation), City University of New York. 231
- Gay, T. J. (1968). Effect of speaking rate on diphthong formant movements. *The Journal of the Acoustical Society of America*, 44(6), 1570–1573.
- Gay, T. J. (1970). *A Perceptual Study of American English Diphthongs*. *Language and Speech*, 3(2), 65–88.
- Gottfried, M., Miller, J. D., & Meyer, D. J. (1993). Three approaches to the classification of American English diphthongs. *Journal of phonetics*, 21(3), 205-229.
- Hanson, H. M. (1995). Glottal characteristics of female speakers—acoustic, physiological, and perceptual correlates. *The Journal of the Acoustical Society of America*, 97(5), 3422-3422.
- Harrington, J., and Cassidy, S. (1994). “Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English,” *Lang Speech* 37, 357–373.
- Hewlett, N., & Beck, J. M. (2013). *An introduction to the science of phonetics*. Routledge.
- Holbrook, A., & Fairbanks, G. (1962). Diphthong formants and their movements. *Journal of Speech and Hearing Research*, 5(1), 38–58.
- Hu, F., & Ge, C. (2016). Vowels and Diphthongs in Cangnan Southern Min Chinese Dialect. In *INTERSPEECH* (pp. 978-982).

- Hung, T. T. (2000). Towards a phonology of Hong Kong English. *World Englishes*, 19(3), 337-356.
- Jacewicz, E., Fujimura, O., & Fox, R. A. (2003). Dynamics in diphthong perception. In *Proceedings of the XVth International Congress of Phonetic Sciences*, Barcelona, Spain (pp. 993-996).
- Jha, S. K. (1985). Acoustic analysis of the Maithili diphthongs. *Journal of Phonetics*, 13(1), 107-115.
- Kachru, B. B. (1982). The bilingual's linguistic repertoire. In *Issues in international bilingual education* (pp. 25-52). Springer, Boston, MA.
- Kachru, B. B. (1983). *The indianization of English: the English language in India*. Oxford University Press.
- Kachru, B. B. (1986). *The alchemy of English: The spread, functions, and models of non-native Englishes*. University of Illinois Press.
- Kachru, B. B. (1992). World Englishes: Approaches, issues and resources. *Language teaching*, 25(1), 1-14.
- Kachru, B. B. (2005). *Asian Englishes: beyond the canon* (Vol. 1). Hong Kong University Press.
- Kachru, B. B. (Ed.). (1992). *The other tongue: English across cultures*. University of Illinois press.
- Kaiser, Ed. (1997) Diphthongs as moving vowels. June 28, 2006. <http://speech.bme.ogi.edu/tutordemos/SpectrogramReading/cse551html/cse551/node28.html>
- Kirkpatrick, Andy (2007) *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.
- Kitagawa, A. (2012). Acoustic characteristics of English diphthongs produced by Japanese learners. In *Proceedings of the 17th Conference of Pan-Pacific Association of Applied Linguistics* (pp. 99-100).

- Kwon, H. B. (2010). Gender difference in speech intelligibility using speech intelligibility tests and acoustic analyses. *The journal of advanced prosthodontics*, 2(3), 71-76.
- Ladefoged, P. (2006). *A Course in Phonetics*. Boston: Thomson Wadsworth.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Ladefoged, P., & Johnson, K. (2011). *A Course in Phonetics 6th Edition*.
- Lass, R. (1984). Vowel System Universals and typology: Prologue to Theory. *Phonology Yearbook*, 1, 75–111.
- Lehiste, I. (1964). *Acoustical Characteristics of Selected English Consonants*. Bloomington: Indiana University.
- Lehiste, I., & Peterson, G. E. (1961). Transitions, glides, and diphthongs. *The Journal of the Acoustical Society of America*, 33(3), 268–277.
- Lindau, M., Norlin, K., & Svantesson, J. O. (1985). Cross-linguistic differences in diphthongs. *The Journal of the Acoustical Society of America*, 77(S1), S100-S100.
- Lindau, M., Norlin, K., & Svantesson, J.-O. (1990). Some cross-linguistic differences in diphthongs. *Journal of the International Phonetic Association*.
- Liu, H., & Ng, M. L. (2009). Formant characteristics of vowels produced by Mandarin esophageal speakers. *Journal of voice*, 23(2), 255-260.
- Low, E. L. (2010). The acoustic reality of the Kachruvian circles: A rhythmic perspective. *World Englishes*, 29(3), 394-405.
- Malmberg, B., & Kaiser, L. (Eds.). (1968). *Manual of phonetics*. North-Holland.
- Mayan, D. (1987). *The Indianization of English. The English language in India*.
- McArthur, Tom (1987) *The English languages? English Today*.
- McCully, C. (2009). *The sound structure of English: An introduction*. Cambridge University Press.

- McMahon, A. (2002). *An introduction to English phonology*. Edinburgh University Press.
- Moag, Rodney (1992) *The life cycle of non-native Englishes: a case study*. In Braj Kachru (ed.) *The Other Tongue: English Across Cultures*, 2nd edn. Urbana, IL: University of Illinois Press.
- Nearey, T. M., and Assmann, P. F. 1986. "Modeling the role of inherent spectral change in vowel identification," *J. Acoust. Soc. Am.* 80, 1297– 1308.
- Norlin, K. (1984). *Acoustic Analysis of Vowels and Diphthongs in Cairo Arabic*. Working papers/Lund University, Department of Linguistics and Phonetics, 27.
- O. Elyan, "Sex differences in speech style," *Women Speaking*, vol. 4, pp. 4–8, 1978.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the acoustical society of America*, 24(2), 175-184.
- Pickett, H. M., Poynter, R. L., Cohen, E. A., Delitsky, M. L., Pearson, J. C., & Müller, H. S. P. (1998). Submillimeter, millimeter, and microwave spectral line catalog. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 60(5), 883-890.
- Pike, K. L. (1947). On the phonemic status of English diphthongs. *Language*, 23(2), 151–159. Press.
- Ren, H. (1986). *On the acoustic structure of diphthongal syllables*. University of California, Los Angeles.
- Schneider, Edgar (2003) *The dynamics of new Englishes: from identity construction to dialect birth*. *Language*.
- Schneider, Edgar (2007) *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.
- Simpson, A. P., & Ericsson, C. (2003, August). Sex-specific durational differences in English and Swedish. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1113-1116).
- Stevens, Peter (1980) *Teaching English as an International Language: From Practice to Principle*. Oxford: Pergamon.

Sweet, H. (1890). A primer of phonetics. Clarendon Press.

Trager, G. L., & Smith, H. L. J. (1951). An Outline of English Structure. Norman: Battenburg Press.

Wouters, J., & Macon, M. W. (2002). Effects of prosodic factors on spectral dynamics. I. Analysis. *The Journal of the Acoustical Society of America*, 111(1), 417-427.

Yin, L. (2014). An instrumental comparison of diphthongs in Malaysian English and Singapore English (Doctoral dissertation, University of Malaya Kuala Lumpur).

Appendix-A

The following tables (appendix-A & B) show F1, F2 (in hertz) and the total duration (in milliseconds) of each diphthong pronounced 4 times embedded in 4 different English words. As for the range of frequency, the standard is 0 Hz at the bottom and 5000 Hz at the top. The first column in the following tables lists the token words pronounced by eight study participants. The second and third columns show the F1 and F2 values against each target vowel. The fourth column provides the duration of the target vowel. The columns from 5th to 8th provide the same for different vowel sounds. However, the values for / aʊ / are put from columns 1 to 5 for their even distribution.

F1 and F2 values of the selected diphthongs extracted with the help of Praat are being presented in an informative and useful method offered and used by Hung (2000).

F1, F2 & duration - Female participants

Item	F1(HZ)	F2(HZ)	Duration (ms)	Item	F1(HZ)	F2(HZ)	Duration (ms)
Dale	447	1938	299	Hoity	670	1676	186
Grey	476	2189	236	Toity	745	1754	182
Apron	428	1764	215	Boy	692	1835	251
Sale	558	2092	266	Enjoy	703	1736	218
Dale	478	2027	235	Hoity	645	1568	140
Grey	459	2116	190	Toity	647	1710	138
Apron	465	2039	158	Boy	660	1766	244
Sale	498	2071	297	Enjoy	659	1983	235
Dale	429	2418	300	Hoity	592	1961	163
Grey	436	2369	229	Toity	618	1496	201
Apron	458	2192	276	Boy	662	1883	287
Sale	431	2389	248	Enjoy	639	1531	267
Dale	512	2442	230	Hoity	558	1461	168

An Acoustic Investigation into Diphthongs . . .

Grey	548	2280	178	Toity	442	990	163
Apron	447	2645	145	Boy	531	1899	321
Sale	457	2438	207	Enjoy	543	1789	267
Giles	824	1832	367	Old	477	1221	277
White	763	1913	265	Spoke	549	1252	214
Tie	778	2093	261	Croak	470	1152	232
Nice	844	2124	231	Throat	552	1096	144
Gile	564	2041	331	Old	421	1272	241
White	814	1811	240	Spoke	466	1222	237
Tie	751	2021	230	Croak	402	1133	202
Nice	812	1993	218	Throat	444	1065	211
Gile	762	2106	296	Old	472	1182	297
White	961	1683	209	Spoke	492	1351	214
Tie	794	2029	222	Croak	435	1049	248
Nice	885	1780	181	Throat	489	1094	207
Gile	744	1991	326	Old	530	1325	291
White	710	1942	247	Spoke	508	1126	133
Tie	770	1992	291	Croak	451	1187	272
Nice	553	2316	183	Throat	539	1525	212
Stout	740	1290	235	Stout	708	1219	227
Shout	650	1397	195	Shout	763	1352	159
Bout	639	1483	354	Bout	754	1285	216
Round	780	1920	324	Round	802	1351	227
Stout	663	1300	172	Stout	741	1586	217
Shout	834	1464	147	Shout	643	1681	187
Bout	676	1349	180	Bout	718	1705	207
Round	774	1397	176	Round	677	1568	291

d

Appendix-B

F1, F2 & duration - Male participants

Item	F1(HZ)	F2(HZ)	Duration (ms)	Item	F1(HZ)	F2(HZ)	Duration (ms)
Dale	409	1533	206	Hoity	481	1536	153
Grey	440	1623	236	Toity	599	1581	123
Apron	434	1843	135	Boy	477	1704	210
Sale	432	1525	233	Enjoy	480	1716	201
Dale	498	1503	281	Hoity	457	1508	131
Grey	447	2100	237	Toity	489	1656	142
Apron	417	1802	146	Boy	490	1688	397
Sale	413	1685	202	Enjoy	439	1480	310
Dale	404	1614	289	Hoity	422	1364	142
Grey	441	1867	181	Toity	426	1145	113
Apron	371	2182	096	Boy	455	1191	241
Sale	430	1549	258	Enjoy	472	1175	149
Dale	263	2206	240	Hoity	317	1520	159
Grey	246	2041	126	Toity	385	1200	170
Apron	425	1717	106	Boy	471	1271	251
Sale	302	1911	227	Enjoy	134	1803	295
Giles	593	1363	267	Old	418	1442	214
White	520	1478	184	Spoke	470	1111	120
Tie	675	1650	172	Croak	407	1214	157
Nice	644	1425	216	Throat	459	1478	187
Gile	607	1330	223	Old	427	1070	206
White	539	2031	333	Spoke	444	1136	171

An Acoustic Investigation into Diphthongs . . .

Tie	645	1693	333	Croak	489	1206	263
Nice	612	1698	269	Throat	466	1100	237
Gile	669	1388	245	Old	426	1161	148
White	590	1140	151	Spoke	470	978	143
Tie	571	1468	223	Croak	520	1276	150
Nice	570	1317	232	Throat	467	1135	218
Gile	460	1369	297	Old	479	1016	241
White	395	1151	170	Spoke	466	1330	212
Tie	521	1614	246	Croak	486	1202	126
Nice	256	1399	147	Throat	490	999	151
Stout	623	2094	176	Stout	632	1454	216
Shout	723	2143	146	Shout	620	1174	179
Bout	616	1177	202	Bout	617	1255	212
Round	700	1566	210	Round	532	1406	212
Stout	724	1698	193	Stout	564	1950	175
Shout	775	1805	180	Shout	472	1919	234
Bout	746	1210	210	Bout	558	1874	185
Round	696	1384	268	Round	457	1234	239
