English Problematic Diphthongs for Pashto Speakers: An Acoustic Analysis

Dr. Ghani Rahman
Assistant Professor, English Department, Hazara University Mansehra Pakistan
ghani.rahman@hu.edu.pk

Dr. Abdul Hamid
Lecturer in English, University of Swat Pakistan
abdulhamid@uswat.edu.pk

Dr. Rooh Ul Amin
Associate Professor, Department of English,
University of Sialkot

Mr. Muzaffar Ali
M. Phil Scholar, Hazara University Mansehra

Abstract
Concerned with acoustic analysis of English diphthongs, the data for the present study were collected from undergraduate level Pashtun students of English and were analyzed through PRAAT software. The formants frequencies of F1 and F2 were calculated for these diphthongs pronounced by Pashto speakers at different possible positions of the word. The results show that except three which are, /ɔɪ/, /əʊ/ and /aʊ/,all English diphthongs were problematic for the Pashtun learners of English at undergraduate level. Amongthe problematic diphthongs, some were highly problematic and some were less problematic. Diphthongs such as /ɪə/, /eə/ and /aɪ/ were less problematic. On the other hand, the diphthongs /ʊə/ and /eɪ/ were found highly problematic for the students. The present study concludes that the English diphthongs should be taught with special focus on their pronunciation in English words, particularly, the problematic diphthongs. The study is has significance for English language teachers and learners and syllabus designer to know about English problematic diphthongs for second language learners.

Keywords: English diphthong, articulation, acoustic analysis, Pashto speakers.

Introduction

Language is a unique human fundamental faculty differentiating them from all other animals (Aitchison, 2008). All languages have some subsystems, i.e., phonetics and phonology, morphology and syntax, etc. where transfer from mother tongue into the target language (TL) can be observed. However, this transfer is commonly less morphological and syntactic than phonetic and phonological (Rahman, 2009). The learners' performance of the target language is affected by the similarities and differences of their L1 and L2 (Slabakova, 2016).

Correct pronunciation is an important skill in second language learning (Gilakjani, 2011). The way of speaking a second language (accent) that characteristically change does not only refer to phonetic demonstration but also includes prosodic changes (Mixdorff, Cole & Shattuk-Hufnagel, 2012) and so automatic speech recognition systems are mostly affected by accent (Benzeghiba et al., 2006). This study explores English problematic diphthongs for Pashto speakers by acoustically analyzing diphthongs to establish the most problematic diphthongs. The acoustic values of these diphthongs are compared with acoustic values in Standard Southern English.

Statement of the Problem

Second language learning is difficult in several ways. Pronunciation is one the most important aspect of these problems (Dejica-Cartis & Dejica-Cartis, 2013), which may be because of L1 and L2 structural differences (Bell 1995) and most of the errors committed are because being unaware of the structures (speech sounds) of the target language (Aljomaely, 2018). English being the target language has greater problems for second language learners in terms of speech sounds of English. Out of these speech sounds, the vowels either short, long and diphthongs are especially problematic for Pashto speakers. Out of all these vowels, diphthongs are of special importance because they are a combination of phonetically two short vowels. A diphthong is a combination of two vowels, for the production of which the tongue glides from one position to the other. This movement of the tongue from one position to another make the diphthongs difficult to learn and use. Secondly, every language has different number and types of diphthongs. The number and quality of Pashto diphthongs are different from that of English; similarly, the English diphthongs are different from that of Pashto. When Pashto speakers pronounce English diphthongs, they pronounce them like Pashto diphthongs. The present study with the help of acoustic analysis (by using PRAAT software) identifies the number of these English problematic diphthongs. The number and nature of these English problematic diphthongs have not been identified in any previous studies, so, the present study focuses on these problems in order to investigate English problematic diphthongs for Pashto speakers.

Research Objectives

The present study has the following objectives:

1. To identify English problematic diphthongs for Pashto speakers.
2. To find out the nature of these problematic diphthongs for Pashto speakers.

Review of Related Literature

Most of the studies (e.g., Aljomaely, 2018; Slabakova, 2016) have focused on the discrepancy between English letters and speech sounds. The English spelling system, particularly, makes this more problematic (Roach, 1991). These phonetic/phonological features of speech sound can only be studied scientifically through adequate and reliable tools, though these tools too, have their own limitations and inherent shortcoming (Furui, 2005) but still they are better be used to make study more scientific in nature. The contextual change generally and environmental change particularly, are challenging for speech recognizers (Anusuya, 2009). The following three approaches are the most commonly used by speech recognizers:

a) Pattern recognition
b) Artificial intelligence
c) Acoustic phonetics

Out of these approaches, the acoustic approach divides, labels and locates the distinct phonetics elements of signals of speech (Espy-Wilson, 1987, 1994). A host of research has focused on the acoustic phonetics, the vowels sounds being focused more in these studies (Kent & Charles, 2002).

English Diphthongs

English language (Standard English) has long and short vowels along with eight diphthongs, in English, which are shown in the following figure. It also shows the tongue position in initial vowel position and its movements toward the preceding vowel position.

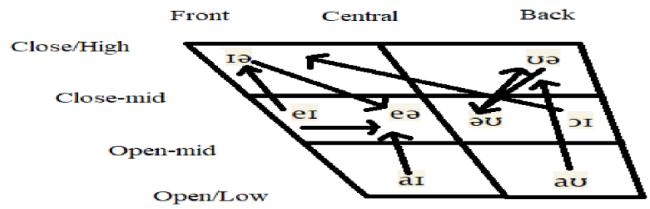

Figure 1: Eight English diphthongs on a vowel grid

A diphthong is phonetically a combination of two short vowels. Diphthong in English is a combination of two short vowels. For example, the vowel sound in /daɪd/ is a diphthong /aɪ/. The word diphthong is originally derived from the word, *diphthongos* (a Greek word)which is a combination of "*di"* and "*phthongoss"* meaning 'twice' (from *di)* and *phthongos* means 'voice, sound'. In English, a diphthong is a vowel sound having special features. There is a deliberate movement (glide) of the tongue from one vowel position to another vowel position, and which is produced in one syllable (Ramelan, 1999). The glide (movement) is the movement for the speech from position to another. Diphthong therefore, like the diphthong /aɪ / is represented by two vowel symbols (in case of English, they are two vowel sounds), though the movement for second vowel is never reached (Ramelan, 1999), since both the sounds are automatically produced. Although a diphthong is phonetically a combination of two short vowels, but both the vowels are not equally syllabic since both are produced at the same time in one syllable which is produced/uttered with a single chest pulse (single puff or air). If a combination of tow vowels cluster is produced with two chest pulses (impulses of breath), the resultant combination is not a diphthong but a dissyllabic word or a common sequence to two vowels. For example, in the word 'die' /daɪ/, the vowel combination /aɪ/ is a diphthong. Out of these two vowels, only one is more sonorous or louder (syllabic) than the other (the other one is non-syllabic) (Ramelan, 1999). English diphthong are either centering diphthong or closing diphthong. In centering diphthong, the second vowel is more center than the first vowel sound; while in closing diphthong, the second vowel is closer. The English centering diphthongs are [eə], [ɪə] and [ʊə]. The English closing diphthongs are [eɪ], [əʊ], [aɪ], [aʊ], and [ɔɪ].

Acoustic Properties of English Diphthongs

The articulation of a diphthong involves a change in quality of one vowel sound to another vowel sound (Ladefoged & Johnson, 2006). Diphthong is produced when two vocalic sounds come together along voicing of vocal cards in the larynx. During such utterances, the tongue position changes and produces one vowel (quality of the vowel) and the other (vowel quality) vowel follows. The size and tract of the vocal cavity varies primarily for a speaker by the positioning of the lips and tongue, thus altering the phonetic quality of vocalic sounds with altering the tract. The position and shape of the tongue along the protrusion of lips define the phonetic qualities of the vowel sound. The tongue also determine the geometry of the pharyngeal and oral cavity controlled by the areas and shape of the front of the vocal tract helps the protrusion of the lips extending the overall length of the tract (Clark & Yallop, 1994). Diphthongs (like other vowels sounds) are analyzed in the form of spectrographic representation for their 1st and 2nd formant frequency levels capturing the shape of the resonant property of the articulatory cavity of the dissimilar vowel sounds (Ball & Rahilly, 2014). The tongue height, tongue advancement, and lip-rounding classify the vowel sounds. The change of the tongue movements is shown in the spectrogram via formant patterns for eight diphthongs. For studying the direction of diphthongal glide as shown in Figure 1 and Figure 2, the direction of each was analyzed.

Previous Works on English Diphthongs

Many studies have been conducted on diphthongs like the study of Fry (1979), Lehiste & Peterson (1961), Holbrook & Fairbanks (1962), Leimgruber (2011), Lee & Lim (2000), Gay (1968), Ladefoged & Johnson (2006), Hayward (2000) and Deterding (1997). Some like Fry (1979) are of the view that English diphthong is the combination of two pure vowel sounds consisting of one syllable. For Hayward (2000), it is a representation of a sequence of two vowel sounds, the first of these two sounds represents the starting point and the second of these sounds represents the ending point. These descriptions were not accepted by many other scholars. For example, according to Ladefoged & Johnson (2006), diphthongs involve movements within a single syllable from one vowel to another vowel. But, in Lehiste's and Peterson's (1961) view, measuring a diphthong is actually the measurement of the duration of on glide from the consonants release to the steady state to the end of the off glide of the vowel. Along the same line, for Gay (1968), ROC (the rate of change) governs a diphthong of the formant transitions rather than target position of the offset or onset.

Many other studies confirmed and supported that the formant rates of change remain constant even when there are considerable variations in the off glide and on glide position (Lee & Lim, 2000; Deterding, 1997; Gay, 1968). The diphthongs are sometimes produced as monophthongs in many varieties of English. This monophthongization was reported in many studies in British English (e.g., Roach, 2010; Foulkes & Docherty, 2007), for American English (e.g., Kiesling & Wisnosky, 2003; Johnstone & Kiesling, 2008), for Australian English (e.g., Trudgill & Hannah, 1985), and for New Zealand English (e.g., Trudgill & Hannah (1985).

Research Methodology

The purpose of this study was to find out English problematic diphthongs and the problems that students face in term of phonological patterns. The research was based on experimental design which comprised quantitative approach for its data analysis. The quantitative approach of speech sounds utilizing experimental techniques and acoustics approach of recorded speech sounds (Li, 2004; Ball & Rahilly, 1999). The data were collected from students whose first language was Pashto speakers had background of English language through scientific tool voice recorder. Non-random convenient sample was used for collecting data from sixteen participants from different territories of Khyber Pakhtunkhwa speaking Yousafzai dialect. The age limit of the participants ranged between twenty and twenty-nine years. The participants were asked to pronounce different words including target diphthongs at three conceivable positions that are, initial, middle, and final position of the word. The recording device had the frequency range of 20 KHz to 40 Hz. High quality Speakers were additionally utilized to examine so as to listen to the articulation of members unmistakably.

The data collected (voice recordings) from Pashto speakers learning English were analyzed through computer generated PRAAT, a software which gives spectrogram for data observation. Consequently, the pronunciation of the target diphthongs was recognized later the examination of the formants, which are F2 and F1 of the data through spectrograms. Various studies have suggested diverse approaches to designate diphthongs acoustically. For example, Ren (1988) measured in detail the F2 trajectory in diphthongal syllables at various points, followed by Maxwell's and Fletcher's (2010) approach presenting the time normalized average formant F1 and F2 trajectory for an individual diphthong at different points. The problematic vowel sounds for Pashto speakers were recognized by observing the recorded sounds. So, the target diphthongs in the were observed and their manner of articulation was analyzed acoustically through spectrograms. Following Deterding, Wong and Kirkpatrick (2008) and Maxwell and Fletcher (2010), the F1 and F2 frequencies for each individual diphthong were used to track its trajectories,. The reading was taken carefully for measurement later avoiding formant transition. The trajectory for each diphthong was linearly inserted and the time was normalized with the average formant frequency plotted for analysis.

Results

The height of the vowel has an inverse proportion to F1 values, which is either high or close or vowels having lower F1 values, than low or open vowel sounds (Ball & Rahilly, 1999). F2 values reflect tongue advancement, where the front vowels have higher F2s, than back vowels (Ball & Rahilly, 1999). Nonetheless, the degree of backness and correlation between the second formant frequency is not as good as the correlation the vowel height and the first formant frequency (Ladefoged & Johnson, 2006). The reason for this is the degree of lip rounding and the vowel height that substantially affects the second formant frequency. lowering of the second and third formants usually characterize lip rounding. Usually, the movement is from the more prominent vowel to the other vowel (Brosnahan & Malmberg, 1970). In diphthongs, there is vowel to vowel glide involving high interdependence between the two vowel sounds. The tongue movement from a vowel to another is not a complete change but it gives rise to rapid switching (more or less) from one set of formants to another (Brosnahan & Malmberg, 1970). In the following sections, the data results about the glide in the form of F1 and F2 for all diphthongs are given.

The English Diphthong /ɪə /

The following spectrogram of target diphthong /ɪə / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /ɪə / at first position of the words.

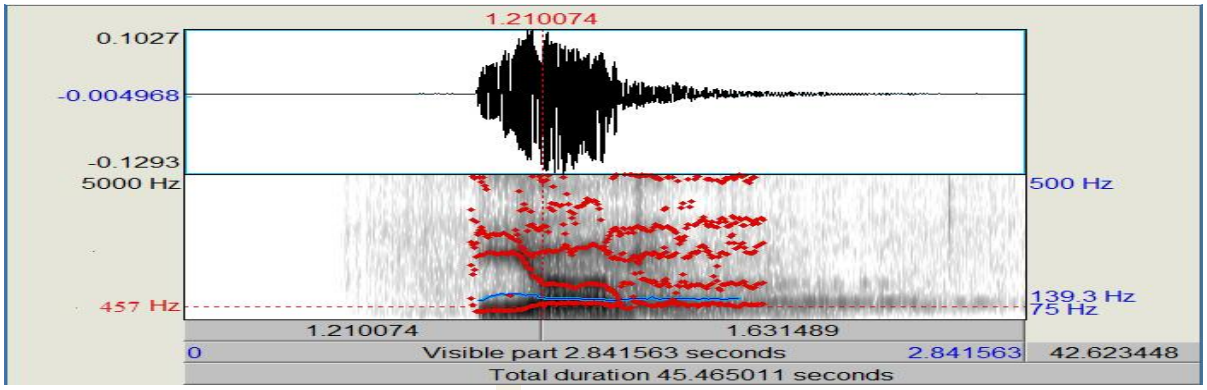

Figure 2: The pronunciation of diphthong /ɪə / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /ɪə /, The columns to left show formant frequency of /ɪə / in Standard English of the target diphthong /ɪə /. The columns to the right show the formant frequency diphthong /ɪə /. The columns to the rightmost side show the mean formant frequency of the target diphthong.

Table 1: Formant frequency of native speakers and Pashto speakers of diphthong /ɪə/

Formant Frequency of /ɪə / by Native Speakers		Formant frequency of /ɪə / by Pashto speakers						Average formant Frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
442	1839	457	1644	493	1428	503	1551	484	1541

The value in the above table shows that the mean formant average frequency value for F1 for diphthong /ɪə/by Pashto speakers was 484 Hz which is nearly the value of F1 442 Hz by native speakers. This value is quite near the standard value at initial position (457 Hz) compared to the value at medial (493 Hz) and final positions (503 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /ɪə / at initial and last position is nearly-native like.Unlike this value, in the above table, the mean formant average value for F2 for diphthong /ɪə/by Pashto speakers was 1541 Hz which was not like the value of F2 1839 Hz by native speakers. This value is different from the standard value at initial (1644 Hz), medial (1428 Hz) and final positions (1551 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /ɪə / is problematic for Pashto speakers both in terms of height and backness of the tongue.

The English Diphthong /eə /

The following spectrograms of target diphthong /eə / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /eə / at first position of the words.

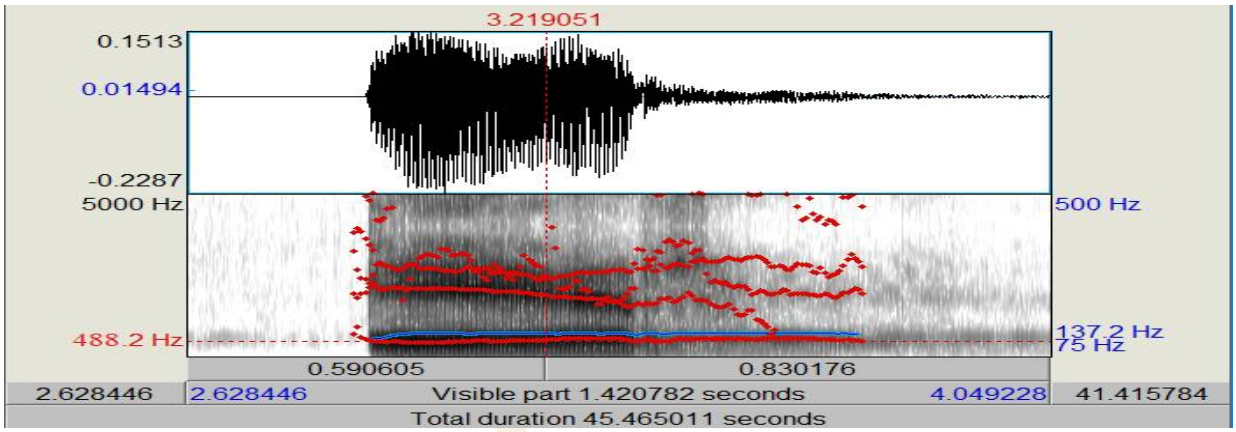

Figure 3: The pronunciation of diphthong /eə / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /eə /. The columns to left show the formant frequency of /eə / in Standard English of the target diphthong /eə /. The columns to the right, show the formant frequency of diphthong /eə /. The columns to the rightmost side show the mean formant frequency of diphthong/eə /.

Table 2: Formant frequency of native speakers and Pashto speakers of diphthong/eə/

Formant Frequency of /eə/ by native Speakers		Formant frequency of /eə / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
542	1656	488	1885	469	1595	485	1822	480	1767

The values, in the above table, show that the mean formant average frequency value for F1 for diphthong /eə /by Pashto speakers was 480 Hz which is a bit different from the value of F1 542 Hz by native speakers. This value is almost the same initial (488 Hz) and final position (485 Hz) compared to the value at medial position (469 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /eə /at initial and final position is a bit different from native like and that is still different from the value at medial position. The result suggest that the tongue is not raised like the raising of the tongue for the same sound by native speakers. This value in the above table for the mean formant average value for F2 for diphthong /eə /by Pashto speakers was 1767 Hz which was not like the value of F2 1656 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1885 Hz) and final position (1822 Hz) which is still different and greater at medial position (1595 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /eə / is problematic for Pashto speakers both in terms of height and backness of the tongue.

The English Diphthong /ʊə/

The following spectrograms of target diphthong /ʊə / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /ʊə / at first position of the words.

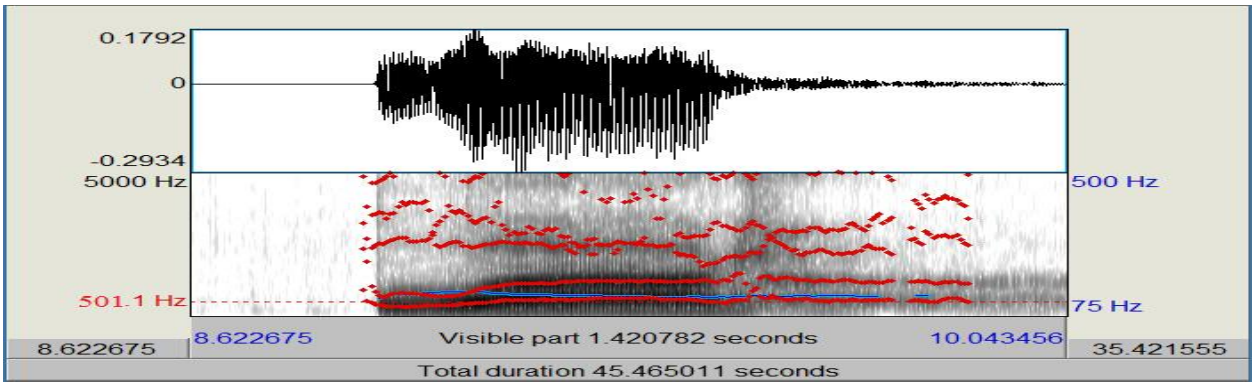

Figure 4: The pronunciation of diphthong /ʊə / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /ʊə /. The columns, to left, show the formant frequency of /ʊə / in Standard English diphthong /ʊə /. The columns to the right show the formant frequency of the target diphthong /ʊə/. The columns to the rightmost side show the mean formant frequency of diphthong.

Table 3: Formant frequency of native speakers and Pashto speakers of diphthong/ʊə/

Formant Frequency of /ʊə/ by native Speakers		Formant frequency of /ʊə / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
607	1565	501	1024	404	1652	517	1127	474	1267

The values, in the above table, show that the mean formant average frequency value for F1 for diphthong /ʊə /by Pashto speakers was 474 Hz which was quite different from the values of F1 607 Hz by natives. The difference was observed in the value of the same diphthong at initial (501 Hz), med position (404 Hz) and last position (517 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /ʊə / at all positions was not native like. The result suggest that the tongue was not raised like the raising of the tongue for the same sound by native speakers. This value in the above table for the mean formant average value for F2 for diphthong /ʊə /by Pashto speakers was 1267 Hz which was not like the value of F2 1565 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1024 Hz) and final position (1127 Hz)while the value at medial position (1652 Hz) is a bit near native like value suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /eə / is problematic for Pashto speakers both in terms of height and backness of the tongue, particularly at initial and final positions.

The English Diphthong /eɪ /

The following spectrograms of target diphthong /eɪ / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /eɪ / at first position of the words.

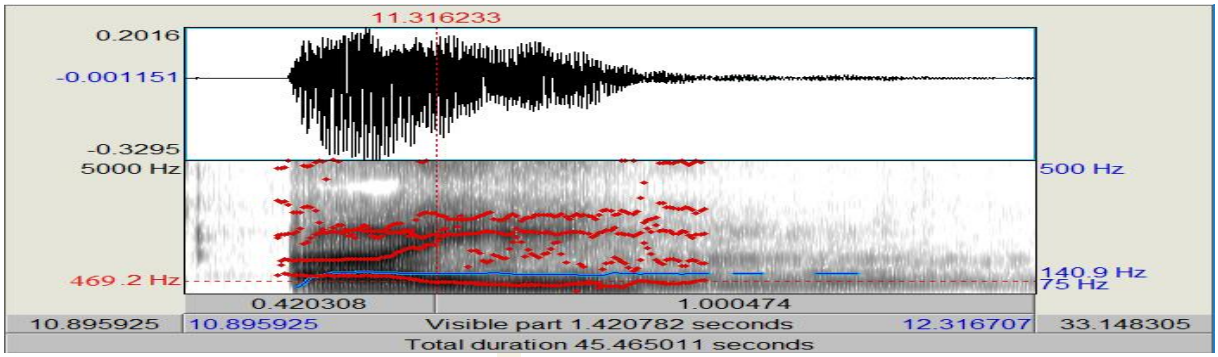

Figure 5: The pronunciation of diphthong /eɪ / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /eɪ /. The columns, to the left, show the formant frequency of /eɪ / in Standard English diphthong /eɪ /. The columns to the right show the formant frequency of diphthong /eɪ /. The columns, to the rightmost side, show the mean formant frequency of diphthong/eɪ /.

Table 4: Formant frequency of native speakers and Pashto speakers of diphthong/eɪ /

Formant Frequency of /eɪ/ by native Speakers		Formant frequency of /eɪ / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
425	1952	469	1815	492	1802	482	1783	481	1800

The values, in the above table, show that the mean formant average frequency value for F1 for diphthong /eɪ /by Pashto speakers was 481 Hz which is a bit greater but not significantly different from the value of F1 425 Hz by native speakers. This value is almost the same at initial (469 Hz), medial (482 Hz) and final position (492 Hz) suggesting tongue's height measured by F1 value, for diphthong /eɪ /was near native like at all positions. The result suggest that the tongue is raised like the raising of the tongue for the same sound by native speakers. But the value in the above table for the mean formant average value for F2 for diphthong /eɪ /by Pashto speakers was 1800 Hz which was not like the value of F2 1952 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1815 Hz) and final (1783 Hz) and medial position (1802 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /eɪ / is problematic for Pashto speakers both in terms of height and backness of the tongue.

The English Diphthong /əʊ /

The following spectrograms of target diphthong /əʊ / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /əʊ / at first position of the words.

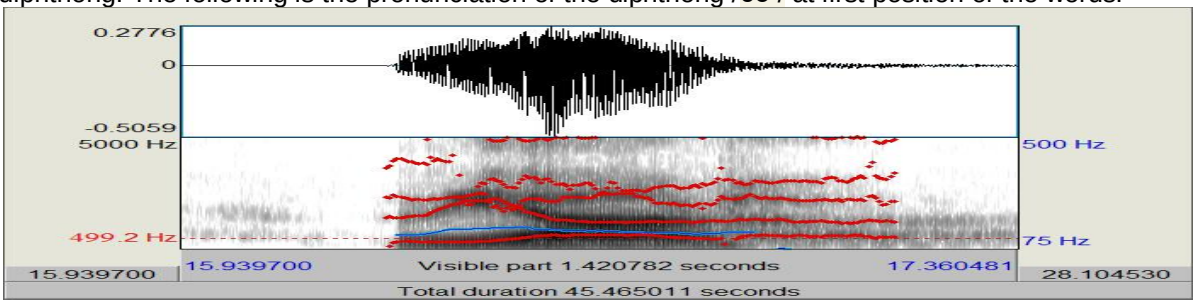

Figure 6: The pronunciation of diphthong /əʊ / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /əʊ /. The columns, to the left, show the formant frequency of /əʊ / in Standard English diphthong /əʊ /. The columns, to the right, show the formant frequency of diphthong /əʊ /. The columns, to the rightmost side, show the mean formant frequency of diphthong/əʊ /.

Table 5: Formant frequency of native speakers and Pashto speakers of diphthong/əʊ /

Formant Frequency of /əʊ / by native Speakers		Formant frequency of /əʊ / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
422	1346	499	1017	497	1081	526	1279	507	1125

The value in the above table shows that the mean formant average frequency value for F1 for diphthong /əʊ /by Pashto speakers was 507 Hz which was a bit different from the value of F1 422 Hz by native speakers. This value is almost the same initial position (499 Hz), final position (526 Hz) and at medial position (497 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /əʊ / at all positions is almost the same like that for native speakers. The results suggest that the tongue raising for this diphthong is nearly the same like for native speakers. This value in the above table for the mean formant average value for F2 for diphthong /əʊ /by Pashto speakers was 1125 Hz which was not like the value of F2 1346 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1017 Hz), final position (1279 Hz) and at medial position (1081 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /əʊ / is problematic for Pashto speakers both in terms of backness of the tongue. Hence we can say that there is no significant difference in terms of height and significant difference in terms of backness for the diphthong /əʊ/.

The English Diphthong /aɪ /

The following spectrograms of target diphthong /aɪ/ shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /aɪ / at first position of the words.

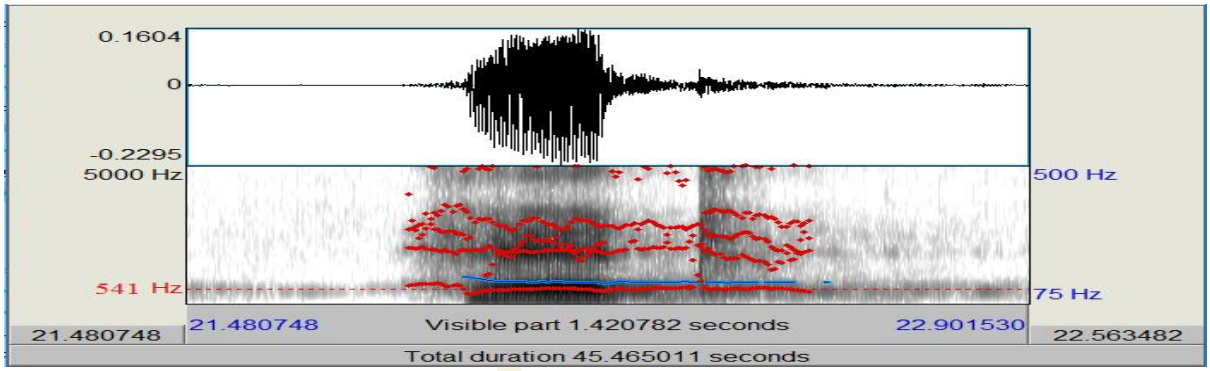

Figure 7: The pronunciation of diphthong /aɪ / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /aɪ /. The columns, on the left, show the formant frequency of /aɪ / in Standard English diphthong /aɪ /. The columns, to the right, show the formant frequency of diphthong /aɪ /. The columns, to the rightmost side, show the mean formant frequency of diphthong/aɪ /.

Table 6: Formant frequency of native speakers and Pashto speakers of diphthong/aɪ /

Formant Frequency of /aɪ / by native Speakers		Formant frequency of /aɪ / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
555	1334	541	1479	526	1416	506	1494	524	1469

The values, in the above table, show that the mean formant average frequency value for F1 for diphthong /aɪ /by Pashto speakers was 524 Hz which is a bit but not significant different from the value of F1 555 Hz by native speakers. This value is almost the same initial (541 Hz), medial position (526 Hz) and at final position (506 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /aɪ / at all positions is almost the same like that for natives. The different is greater at first position and lesser at mid position. The results suggest that the tongue raising for this diphthong is nearly the same like for native speakers. This value in the above table for the mean formant average value for F2 for diphthong /aɪ /by Pashto speakers was 1469 Hz which was not like the value of F2 1334 Hz by native speakers. This value for non-native speakers is different from the standard value at initial position (1479 Hz), final position (1494 Hz) and at medial position (1416 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /aɪ / is problematic for Pashto speakers both in terms of backness of the tongue. Hence we can say that there is no significant difference in terms of height and significant difference in terms of backness for the diphthong /aɪ /.

The English Diphthong /aʊ /

The following spectrograms of target diphthong /aʊ/ show different pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /aʊ / at first position of the words, followed by the pronunciation at medial and final positions respectively.

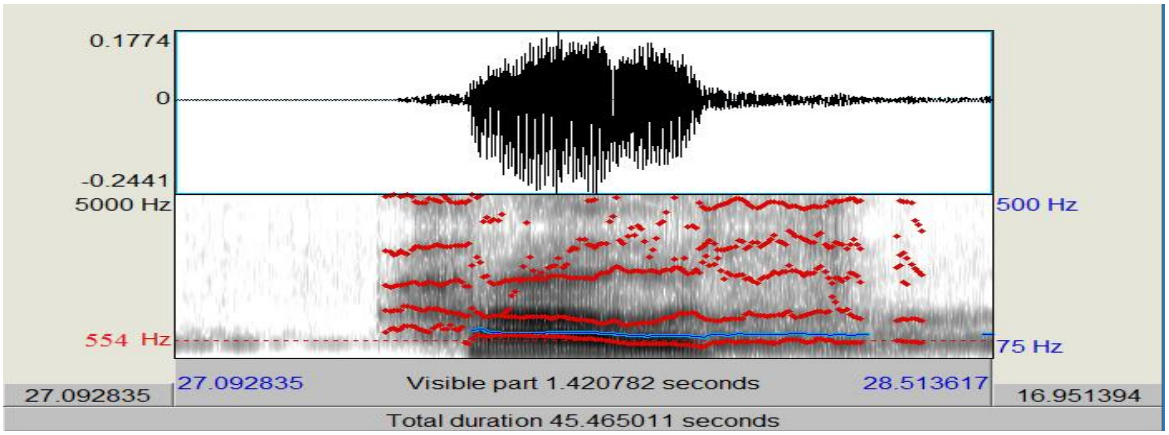

Figure 8: The pronunciation of diphthong /aʊ / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /aʊ /. The columns, to the left, show the formant frequency of /aʊ / in Standard English diphthong /aʊ /. The columns, to the right, show the formant frequency of diphthong /aʊ /. The columns, to the rightmost side, show the mean formant frequency of diphthong/aʊ /.

Table 7: Formant frequency of native speakers and Pashto speakers of diphthong/aʊ /

Formant Frequency of /aʊ / by native Speakers		Formant frequency of /aʊ / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
575	1214	554	1095	570	1169	550	1076	558	1113

The values, in the above table, shows that the mean formant average frequency value for F1 for diphthong /aʊ /by Pashto speakers was 558 Hz which is a bit different from the value of F1 575 Hz by native speakers. This value is almost the same at initial (554 Hz) and final position (550 Hz) compared to the value at medial position (570 Hz) suggesting that the height by tongue measured by F1 value for the target diphthong /aʊ / at initial and final position is a bit different from native like and that is still different from the value at medial position. The results suggest that the tongue is not raised like the raising of the tongue for the same sound by native speakers. This value in the above table for the mean formant average value for F2 for diphthong /aʊ /by Pashto speakers was 1113 Hz which was not like the value of F2 1214 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1095 Hz) and final position (1076 Hz) which is still different and greater at medial position (1169 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /aʊ / is problematic for Pashto speakers both in terms of height and backness of the tongue.

The English Diphthong /ɔɪ/

The following spectrograms of target diphthong /ɔɪ / shows pronunciation of the same target diphthong. The following is the pronunciation of the diphthong /ɔɪ / at first position of the word.

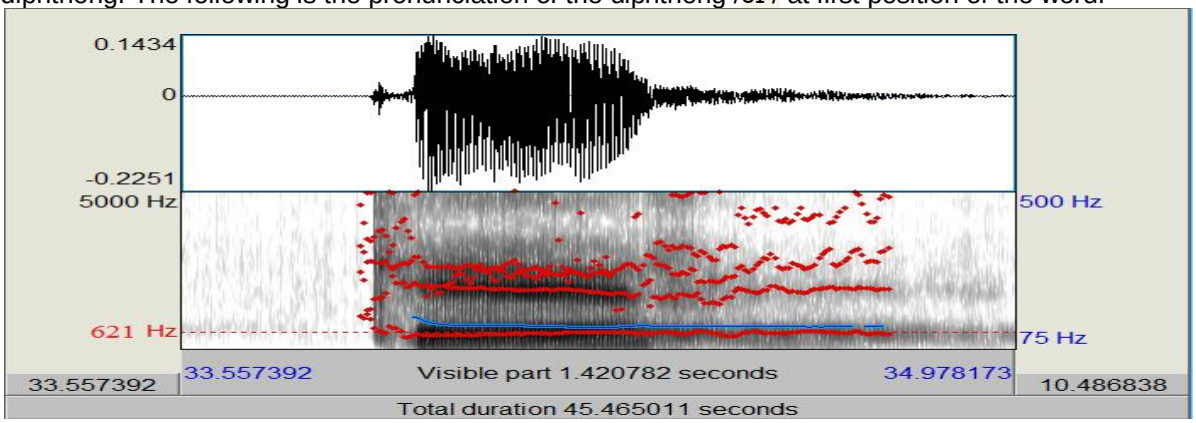

Figure 9: The pronunciation of diphthong /ɔɪ / at first position

The above values acoustically calculated are shown in the following table showing the articulation of diphthong /ɔɪ /. The columns, on the left, show the formant frequency of /ɔɪ / in Standard English of diphthong /ɔɪ /. The columns, to the right, show the formant frequency of diphthong /ɔɪ/. The columns, to the rightmost side, show the mean formant frequency of diphthong/ɔɪ /.

Table 8: Formant frequency of native speakers and Pashto speakers of diphthong/ɔɪ /

Formant Frequency of /ɔɪ / by native Speakers		Formant frequency of /ɔɪ / by Pashto speakers						Average formant frequencies by Pashto Speakers	
		Initial		Medial		Final			
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
461	946	621	1521	509	1553	477	1222	535	1432

The values, in the above table, show that the mean formant average frequency value for F1 for diphthong /ɔɪ /by Pashto speakers was 535 Hz which is a bit different from the value of F1 461 Hz by native speakers. This value is quite different at initial position (621 Hz), compared to its values at final (477 Hz) and medial position (509 Hz) suggesting that the tongue's measured by F1 value of diphthong /ɔɪ / at initial and final position is a bit different from native like and that is still different from the value at medial position. The result suggest that the tongue is not raised like the raising of the tongue for the same sound by native speakers. This value in the above table for the mean formant average value for F2 for diphthong /ɔɪ /by Pashto speakers was 1432 Hz which was not like the value of F2 946 Hz by native speakers. This value for non-native speakers is different from the standard value at initial (1521 Hz) and medial position (1553) which is still different and greater at medial position (1222 Hz) suggesting that the backness of the tongue for this diphthong is not native like. So, the diphthong /ɔɪ / is problematic for Pashto speakers both in terms of height and backness of the tongue.

Discussion

A significant number of mistakes were identified. The mistakes identified in the present study were in terms of some features of English diphthongs and as a result, the participants mispronounced most of the words having English diphthongs. One of the reasons behind this could be phonological differences between the two languages. The results of the data as a whole suggest that Pashto speakers have considerable problems in five English diphthongs, while three diphthongs were not problematic for them. The diphthongs / aʊ /, /aɪ/ and /eɪ /had no problems for them while the diphthongs /ɪə /, /eə /, /ʊə /, /ɔɪ / and /əʊ / were problematic for them. The nature of these problems were different. Pashto speakers do not have problems in diphthongs/ aʊ /, /aɪ/and /eɪ / like investigated by Mlinar (2011) as well. While the diphthongs /ɪə /, /eə / had somewhat problems in Mustikareni's (2013) findings for these sounds with the additional /ʊə/ are also problematic diphthongs in the present study. The diphthongs /ɔɪ / and /əʊ / were highly problematic the same way as described Mustikareni (2013)described. The diphthongs were sometimes confused with other vowel pairs and so affected the pronunciation skills of the students (Martirosian & Davel, 2008). Pronunciation of these diphthongs was often influenced by [consonantal] context in which they appeared (Elvin, Williams, & Escudero, 2016). Almost the same results were found in other studies (e.g., Cebrian, 2019; Din & Rahman, 201; Maxwell & Fletcher, 2010; Tsukada, 2008) where the participants spoke a South Asian language. The variations in diphthongs for these speakers were found because of the factor of mother tongue and the English diphthongs were produced like longer monophthongs (Maxwell & Fletcher, 2010). The movement for diphthongs here was not observed like the movement for native speakers' diphthongs, a property shared by most South Asian languages (Tsukada, 2008). The findings show that English diphthongs were closer to the diphthongs of mother tongue (Din & Rahman, 2011) with combative assimilation to native diphthongs demanding the need for cross-linguistic comparison (Cebrian, 2019).

Conclusion

The results suggest that except three all English diphthongs were problematic for the participants. These problems varied from tongue height to backness of diphthongs. Among the problematic diphthongs, some were highly problematic whereas some were less problematic. Three English diphthongs, which are /ɔɪ/, /əʊ/ and /aʊ/ were not problematic for students. They pronounced these

diphthongs like native speakers. In addition, the diphthongs /ɪə/, /eə/ and /aɪ/ were less problematic, which were pronounced nearly native like. On the other hand, the diphthongs /ʊə/ and /eɪ/ were found highly problematic for the students. Pashto speakers learning English cannot pronounce these diphthongs like native speakers or nearly native like. Most of these problems in the light of the results are claimed to due of lack of any meaningful training in English phonetics and phonology. As a result, the students erroneously pronounced these diphthongs. The pronunciation of these diphthongs might have been influenced by their first language. Even, the correct pronunciation was mostly based on the intuition of the students. So, in learning English diphthongs sounds, correct diphthongs' pronunciation is important. Conscious efforts is required on the part of the learners. Three important factors, i.e., imitations of sounds (leaning correct pronunciation), comparison of speech sounds and practice these sounds is very crucial to be addressed. It is of equal importance that learners should be involved in learning the correct articulation of diphthongs which is only possible if learners are given proper training. If the learners imitate the speech sounds of L2 correctly and compare them with sounds in their L1 and have enough practice of these speech sounds, the correct pronunciation can be improved to a greater extent. This is especially important for English diphthongs, because generally there is one to one correspondence between English letters and English speech sounds. Therefore, the pronunciation of English words should be not learnt from their written forms. To conclude the present study emphasizes that the English diphthongs should be taught with special focus on their pronunciation in English words, particularly, the problematic diphthongs (Mlinar, 2011). Moreover, proper practices of speech sounds are required for learners in order to bring their pronunciation, if not like native speakers, but at least nearer to native speakers' pronunciation.

Recommendations

In the light of the findings of this study, it is recommended that learners (L2 learners) should be properly guided. This becomes especially important in case of speech sounds of L2. The correct pronunciations of English words should be focused in language classes. The present study opens new arenas of exploration and hence, other studies might focus on the relevant items of phonetics and phonology such as short vowels, long vowels, syllables, stress patterns, elision, and assimilation. The researchers can also work on the impacts of regional languages on English sounds and this could be a significant contribution. Finally, the present study also suggests a thorough analysis of Supra-segmental features of English.

References

Aitchison, J. (2008). *The articulate mammal.* Oxon: Routledge.

Aljomaely, R. (2018). The pronunciation of diphthongs by Iraqi learners of English. *Professional Communication and Translation Studies*, 11, 67-72.

Anusuya, M. A. (2009). Speech recognition by machine: A review. *International Journal of Computer Science and Information Security (IJCSIS), 6*(3), 181-205.

Ball, M.J.,& Rahilly, J. (1999). *Phonetics: The science of speech*. New York: Oxford University Press.

Ball, M. J., & Rahilly, J. (2014). *Phonetics: The science of speech*. Routledge.

Bell, J. S. (1995). The relationship between L1 and L2 literacy: Some complicating factors. *TESOL Quarterly, 29*(4), 687-704.

Benzeghiba, M., Mori, D. R., Deroo, O., Dupont, S., Jouvet, D., Fissore, L., Laface, P., Mertins, A., Ris, C., Rose, R., Tyagi, V., & Wellekens, C. (2006). Impact of variabilities on speech recognition. *Multimedia Communications Eurecom,* 7, 3-16.

Brosnahan, L. F., & Malmberg, B. (1970). *Introduction to phonetics*. Cambridge: Cambridge University Press.

Cebrian, J. (2019). Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. *The Journal of the Acoustical Society of America*, *145*(1), 52-58.

Clark, J. Y., & Yallop, C. C. (1994). *An introduction to phonetics and phonology*. Cambridge: Blackwell.

Dejica-Cartis, A., & Dejica-Cartis, D. (2013). Tentative suggestions regarding the development of foreign language oral communication skills in students at universities of applied sciences. *Procedia-Social and Behavioral Sciences*, *84*, 934-938.

Deterding, D. (1997). The formants of monophthong vowels in standard Southern British English pronunciation. *Journal of the International Phonetic Association*, 27, 47-55.

Deterding, D., Wong, J., & Kirkpatrick, A. (2008). The pronunciation of Hong Kong English. *English World-Wide*, *29*(2), 148-175.

Din, R., & Rahman, G. (2011). Acoustic analysis of Pashto vowels. The acoustic analysis of Pashto vowels, *Language in India*, 11 (2) 793-797.

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.

Espy-Wilson, C. (1987). *An acoustic phonetic approach to speech recognition: Application to the semivowels.* (Unpublished doctoral dissertation) Massachusetts Institute of Technology: Massachusetts, United States.

Espy-Wilson, C. (1994). A Feature-based Semivowel Recognition System. *Journal of Acoustic Society of America, 96*, 65–72.

Foulkes, P., & Docherty, G. J. (2007). Phonological variation in the English of England. In D. Britain. (Eds.). *Language in the British Isles*, (pp. 52-74). Cambridge: Cambridge University Press.

Fry, D. B. (1979). *The physics of speech*. Cambridge: Cambridge University Press.

Furui, S. (2005). 50 years of progress in speech and speaker recognition research. *ECTI Transactions on Computer and Information Technology, 1* (2), 64-74.

Gay, T. (1968). Effect of speaking rate on diphthong formant movements. *The Journal of the Acoustical Society of America*, *44*(6), 1570-1573.

Gilakjani, A. P. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education*, 13, 5-7.

Holbrook, A., & Fairbanks, G. (1962). Diphthong formants and their movements. *Journal of Speech and Hearing Research*, *5*(1), 38-58.

Hayward, K. (2000). *Experimental phonetics*. Harlow, UK. Pearson.

Johnstone, B., & Kiesling, S. F. (2008). Indexicality and experience: Exploring the meanings of/aw/-monophthongization in Pittsburgh 1. *Journal of Sociolinguistics*, *12*(1), 5-33.

Kiesling, S. F., & Wisnosky, M. (2003). Competing norms, heritage prestige, and/aw/-monophthongization in Pittsburgh. *Poster Presented at New Ways of Analyzing Variation (NWAV)*, *32*, 9-12.

Kent, R. D., & Charles, R. (2002). *The acoustic analysis of speech*(2nd ed.). San Diego: Singular Publishing.

Ladefoged, P., & Johnson, K. (2006). *A course in phonetics*. Cambridge: Cambridge University Press.

Lee, E.M. and Lim, L. (2000). Diphthongs in Singaporean English: Their realisations across different formality levels, and some attitudes of listeners towards them. In Brown, A., Deterding, D. and Low, E.L. (Eds.), *The English Language in Singapore: Research on Pronunciation* (pp. 100-111). Singapore: Singapore Association for Applied Linguistics.

Lehiste, I., & Peterson, G. E. (1961). Transitions, glides, and diphthongs. *The Journal of the Acoustical Society of America*, *33*(3), 268-277.

Leimgruber, J. R. (2011). Singapore English. *Language and Linguistics Compass*, *5*(1), 47-62.

Levi, V. S. (2004). *The representation of underlying glides: A cross-linguistic study.*(Unpublished doctoral dissertation). University of Washington: Washington.

Martirosian, O., & Davel, M. (2008). Acoustic analysis of diphthongs in standard South African English. *EnglishWorld-wide*, *29*(2), 19-31.

Maxwell, O., & Fletcher, J. (2010). The acoustic characteristics of diphthongs in Indian English. *World Englishes*, *29*(1), 27-44.

Mlinar, R. (2011). *Pronunciation of English diphthongs by speakers of Serbian:Acoustic characteristics* (Unpublished Master's thesis). University of Novi Sad, Serbia.

Mixdorff, H., Cole, j., & Shattuk-Hufnagel, S. (2012*).* Prosodic Similarity-evidence from an Imitation study. *Speech Prosody.* 14, 20-30.

Mustikareni, D. (2013). *Error Analysis on English Diphthongs Pronounced by the Students of SMA Negeri 1 Comal and SMA Negeri 1 Banjarnegara in English Debate on" Relax" TVProgram of TVRI* (Unpublished doctoral dissertation), Universitas Negeri Semarang.

Rahman, G. (2009). *A comparative study of Pashto and English phonology for English language teaching and learning* (Unpublished master's thesis). UMT, Lahore, Pakistan.

Ramelan. (1999). *English phonetics.* Semarang: *CV*. IKIPSemarang Press

Ren, H. (1988). *On the acoustic structure of diphthongal syllable* (Doctoral Dissertation). University of California, Los Angeles.

Roach, P. (1991). *English phonetics and phonology.* Cambridge: Cambridge University Press.

Roach, P. (2010). *English phonetics and phonology: A practical course (4th ed.)*. Cambridge: UK.

Slabakova, R. (2016). *Second language acquisition*. Oxford: Oxford University Press.

Trudgill, P., & Hannah, J. (1985). *International English: A guide to standard varieties of English.* London:Edward Arnold.

Tsukada, K. (2008). An acoustic comparison of English monophthongs and diphthongs produced by Australian and Thai speakers. *English world-wide*, *29*(2), 194-211.