CHAPTER 2

LITERATURE REVIEW

This chapter begins with a discussion of the influence of first language (L1) on second (L2) and foreign language pronunciation in the literature, and then proceeds to examine the similarities and differences between English and Persian vowels. Subsequently, the use of acoustic analysis to examine vowels is discussed. Finally, studies on the production of English vowels in selected varieties of English are reviewed.

2.1 The Influence of L1 on L2 Pronunciation

The major problem that most EFL learners face is their inability to achieve an "acceptable" and "intelligible" pronunciation of English words. While many EFL learners are able to master the different elements of English language like grammar, vocabulary, writing and reading, they grapple with the production of target-like sounds.

Celce-Murcia (1991) posits that mother tongue transfer has more influence on pronunciation than other areas of L2 learning like grammar and vocabulary. As mentioned in the previous chapter, the inability to master the pronunciation of a foreign language results in what is perceived as a "foreign accent" (Tsukada, 1999, p. 373). Rajadurai (2007, p. 91) points out that "[t]he problem is that anyone who does not speak with a native accent [whatever that might be] is stigmatized as speaking with a foreign accent". However, having a foreign accent does not mean that one's pronunciation is "bad", and as Nair, Krishnasamy and de Mello (2006, p. 33) argue, "[s]omeone with a foreign accent can speak the language perfectly intelligibly, and carry on all kinds of communication without hindrance". Thus, the main contention is that a speaker must be understood by listeners, and even a native speaker of English may not be intelligible to someone who is not familiar with the speaker's variety of

English since it is a language that is spoken with a multitude of accents. Many factors affect intelligibility apart from pronunciation such as the "context of use, [which is] a complex setting involving factors related to the speaker, the listener, the linguistic and social context, and the environment" (Pickering, 2006, p. 221). In other words, notions of intelligibility need to go beyond the recognition of sounds but involve the transfer and negotiation of meaning as well. However, "although pronunciation was by no means the sole cause of … communication breakdown, it was by far the most frequent and the most difficult to resolve" (Jenkins 2002, p. 87).

The influence of L1 on L2 pronunciation was supported within the Contrastive Analysis Hypothesis (CAH) proposed by Lado (1957) and by studies such as those by Suter (1976) and Hammerly (1982). Whilst the main assumption underlying CAH was that differences in the L1 and L2 sound system were likely to cause difficulties in producing native-like pronunciation, other theories emerged suggesting that similarities in phonetic categories could be a source of potential difficulty (e.g. Flege, 1987). The basis of this theory is that L2 sounds that are perceived to be “similar” to the ones in L1 will be deemed to be the same, and therefore, learners will not develop new phonetic categories for them, resulting in them being produced in the same way as they are in the L1. However, the exact way in which perceived similarity between L1 and L2 sounds operates remains unclear.

Findings from Tsukada (2008) on the production of English vowels by Thai speakers showed that in terms of vowel quality, Thai speakers produced four English monophthongs /ɪ/, /æ/, /ʊ/, /ʌ/ and two diphthongs /eɪ/, /oʊ/ similar to native speakers of Australian English speakers. However, these four monophthongs were produced shorter than Australian English vowels while the diphthongs were longer than those produced by Australian speakers produced. Thai speakers produced English diphthongs /eɪ/, /oʊ/ equivalent with the long vowels /eː/ and /oː/ in their first language (Thai). According to Tsukada (2008, p. 206), the reason for this is that "the Thai speakers equate English monophthongs and diphthongs with short and long Thai vowels, respectively, and have exaggerated a durational difference between the two classes of English vowels". Another study by Tsukada (2009) showed that since Thai have phonemic vowel length contrast, they have no problems with length contrast in their production of English vowels. However, Thai speakers produced shorter /ɪ/ and longer /i/ than Australian speakers. This occurred because the Thai speakers transferred phonemic length contrast in their first language to the production of their English vowels, resulting in them sounding different from the native speakers of English. On the other hand, Tsukada (2008) found that Thai speakers were able to produce English vowels that were non-existent in Thai language.

These findings confirm Flege's Speech Learning Model (SLM) that phonemes with higher levels of dissimilarities are easier to learn than those which are similar because 'differences' are more noticeable and learners can create a new strategy for particular phoneme (Flege, 1995). Findings from a Malaysian study (Yap, Wong, & Adi Yasran, 2010) demonstrate the relationship between perception and production. The Malay-English bilingual informants in this study generally found it difficult to discriminate between the front vowel pairs of /ɪ/-/i/, and /ɛ/-/æ/ probably due to the lack of quality and length contrast in Malay. Thus, it can be expected that such speakers will conflate these vowel pairs in their production and this is indeed shown in Pillai, Mohd. Don, Knowles, and Tang (2010) and Tan and Low (2010).

However, the whole notion of what exactly constitutes similar properties is unclear. Studies like Munro (1993) found that the vowels that were dissimilar in Arabic and English were not produced native-like by Arabic speakers as predicted by Flege's theory. Munroe's study was conducted on the production of English vowels by Arabic speakers from countries like Kuwait, Jordan, Sudan, Saudi Arabia, Syria, Palestine and Egypt where English is considered as a foreign language. The findings were compared with 23 native speakers of American English. The vowels produced by Arab speakers displayed different characteristics from the native speakers. Munroe's findings showed the larger ratios of short to long vowels durations, reduction of vowel length, and reduction of diphthongs compared to American English.

Similar findings were reported by Iverson and Evans (2007) who examined the production of English vowels among French, Spanish, German, and Norwegian speakers. Their findings indicated that the Spanish and Norwegian speakers of English did not show evidence of having learnt English /əʊ/ that has more dissimilarity to their first languages while, the German speakers showed the ability of learning /ɑɪ/ which has a higher rating of assimilation. Iverson and Evans (2007) derived that "speakers from different L1s learned new aspects of the English vowel system rather than simply assimilating vowels into existing first language categories" (p. 2842).

Nevertheless, the influence of L1 on L2 pronunciation continues to emerge, for example, within the Optimality Theory framework (Eckman, 2004), and therefore a comparison of the acoustic properties of the vowels in L1 and L2 may be able to provide insights into potential learning and production difficulties.

2.2 English and Persian Vowels

Generally, Persian and English vowels differ in three areas: number of vowels, distinctions between vowel quality, and vowel length contrast. The following sections will compare the vowels in Persian and English in order to identify potential areas of difficulty in the production of the latter by Persian speakers.

2.2.1 Monophthongs

A monophthong is considered as a pure vowel with unchanging articulation from its onset to offset. There are considered to be 12 monophthongs in Standard British English generally referred to as Received Pronunciation (RP): /iː/, /ɪ/, /e/, /æ/, /ʌ/, /ɑː/, /ɜː/, /ə/, /ɒ/, /ɔː/, /ʊ/, /uː/ (Gut, 2009, p. 63). However, other varieties of English may have fewer monophthongs (Chen, Robb, Gilbert, & Lerman, 2001). For instance, General American English (GA) has approximately eleven monophthongs (/i/, /ɪ/, /ɛ/, /æ, /ɑ/, /ɔ/, /ʊ/, /u/, /ʌ/, /ɝ/, /ə/) (Ladefoged, 2006, p. 39) since /ɒ/ tend to be realised as /ɑ/, for example in the word *pot* BrE /pɒt/ and AmE /pɑːt/. Both Ladefoged (2001a, p. 28) and Roach (2000, p. 36-38) list 11 monophthong vowels in Standard British English (BBC English) with /ə/ being omitted. This is because while /ə/ is the most frequently used monophthong in English (Cruttenden, 1994, p. 137), it is generally used in unstressed syllables, and does not contrast with /ɜː/. In some descriptions of Standard British English generally referred to as Received Pronunciation (RP), the symbol /ɛ/, which is more open than /e/ is used. Even Ladefoged uses two different symbols: /e/ in 2001a and /ɛ/ in 2006, while Gut (2009, p. 63), Knowles (1987, p.50), and Roach (2000, p.16) use /e/ for RP. Instrumental studies on RP vowels indicate that this vowel is becoming more open, and hence the use of /ɛ/ instead of /e/ (e.g. Hawkins & Midgley, 2005).

Vowel quality is another term for categorising vowels which makes one vowel sound different from other sounds, for example, the difference between words *barney* /ɑː/ and *bunny* /ʌ/. The quality of a vowel depends on the position of tongue, lips and jaw. Gimson (1970) uses the term close, half-close, half open and open while Ladefoged (2006) uses the term high, mid-high, mid, mid-low and low which essentially refer to the degree of jaw opening. The classification of monophthongs depends on several parameters: tongue movement, height of tongue, degree of retraction of tongue and shape of lips. Figure 2.1 illustrates the classification of RP British English monophthongs according to vowel fronting/retraction and vowel height.

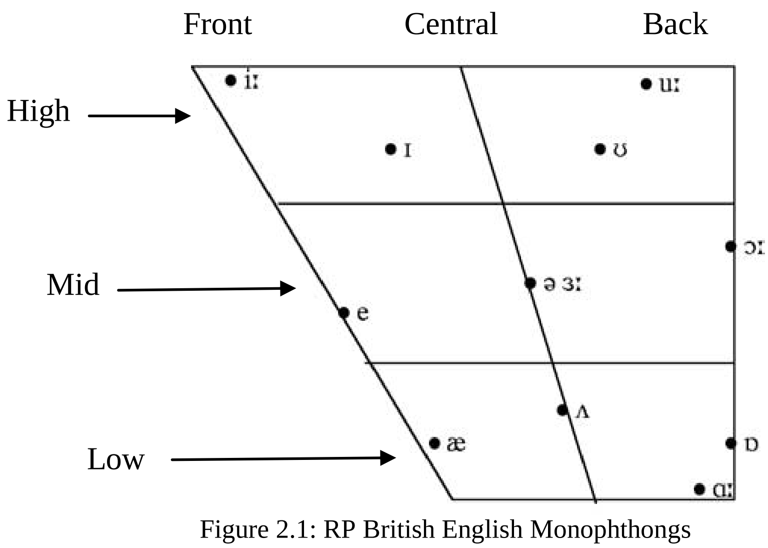

Figure 2.1: RP British English Monophthongs
(Source: Roach, 2000, p. 36-38)

As illustrated in the vowel chart in Figure 2.1, the English vowels, /æ/, /ʌ/, /ɒ/, /ɑː/ can be classified low vowels, while /e/, /ɜː/, /ə/, /ɔː/ as mid and /iː/, /ɪ/, /uː/, /ʊ/ as high vowels. The vowels /iː/, /ɪ/, /e/, /æ/ can be classified as front vowels, /ə/, /ɜː/, /ʌ/ as central and /uː/, /ʊ/, /ɔː/, /ɒ/, /ɑː/ as back vowels.

In comparison to English, there are six monophthongs in Persian: /i/, /u/, /ɒ/, /æ/, /e/, /o/. According to most studies (e.g. Najafi, 2001; Samareh, 2000; Windfuhr, 1979), the three vowels /i/, /u/, and /ɒ/ are considered as long vowels while /æ/, /e/, and /o/ are short vowels. Windfuhr (1979, p. 526) considers vowel length as an active feature however, length is not contrastive in Persian, although the duration of vowels tends to be longer in stressed syllables similar to English (Bashiri, 1991; Keshavarz, 2001). The Persian monophthong vowels are shown in Table 2.1.

Table 2.1: Persian Monophthong Vowels

	Front	Centre	Back
High	i		u
Mid	e		o
Low	æ		ɒ

(Source: Jabbari, 2005, p. 65)

The Persian monophthongs /e/, /u/, /æ/, /o/, /i/ are similar to English vowels. However, Persian language does not have /ɑ/, /ʊ/, /ɪ/, /ɜː/, /ʌ/, /ə/. The comparison between English and Persian vowels (Yavas, 2006, p.197) is shown in Figure 2.2.

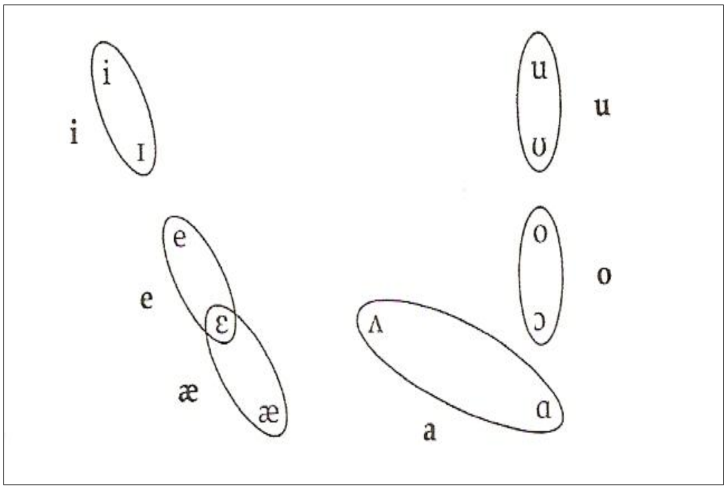

Figure 2.2: Comparisons between English vowels (in ellipsis) and Persian Vowels (Source: Yavas, 2006, p. 197)

As illustrated in Figure 2.2, in Persian /i/ is similar to the close-front /i/ in English but /ɪ/, which is a half-close-front vowel in English, does not exist in Persian. The English monophthong /æ/ which is a low-front vowel has an equivalent vowel in Persian. For example, *sar* /sær/ (*head*) (Jabbari, 2005, p. 71). Hall (2007, p. 38) in her research about the phonological characteristics of Persian speakers of English explains that "Persian /e/ sound corresponds to the English vowels /ɜ/ and /e/ depending on whether it is in either a stressed or an unstressed position". Depending on which theoretical perspective one takes, either the vowels that are not present in Persian and are categorically different from the ones in English or the ones that are similar in both Persian and English will pose problems for Persian EFL learners (see Figure 2.1).

2.2.1.1 Vowel Contrasts

As discussed in the previous section, English vowels differ in quality in relation to the position of the tongue, lips and the jaw, but the vowels can also be categorised as short and long vowels. With regards to length, English monophthongs are divided into two groups; seven short vowels /ɪ/, /e/, /ɒ/, /ʊ/, /ʌ/, /æ/, /ə/ and five long vowels /iː/, /ɜː/, /ɑː/, /uː/, /ɔː/. The duration of sustaining a vowel can be used to distinguish, for example, /ɪ/ from /iː/, /ʌ/ from /ɑː/, /ʊ/ from /uː/, and /ɒ/ from /ɔː/. However, the co-articulatory effects of neighbouring sounds and syllable type can affect vowel length. As Cruttenden (1994, p. 89) asserts "length is not a constant distinctive feature of the vowel for it depends on the context …". For instance, vowels in open syllables tend to be longer than when there are in closed syllables (e.g. *seed* versus *see*). Further, voiceless consonants tend to shorten preceding vowels (/ɑː/ in *cart* is likely to be shorter in duration than in *card*). Vowel duration is also longer before sonorants than before obstruents (e.g. *hurl* compared to *herd*) (Gut, 2009, p. 64). Because of these factors, Gut points out that "phonemically 'long' vowels can actually be shorter than phonemically 'short' vowels when their real duration is measured" (Gut, 2009, p. 64-5).

Long and short vowels can also be linked to "the degree of muscular activity involved in the articulation and to the length of the vowels in question" (Shriberg and Kent, 2003, cited in Bauman-Wängler, 2004, p. 18); Tense vowels compared to lax vowels have more muscle activity and are longer duration as well. For example, Giegerich (1992, p. 99) exemplifies this concept with the vowels /iː/ in the word *beat* /biːt/ which is characterised as tense/long and /ɪ/ in the word *bit* /bɪt/ which is regarded as lax/short. However, Ladefoged (2001b, p. 80-82) explains that these tense and lax categories are "just labels that are used to distinguish two groups of vowels that behave differently in English words" (p. 80). He also defines that tense and lax vowels differently based on their position in a syllable. For instance, tense vowels can occur in open syllables, (e.g. in *see* and *do*) while, the lax vowels (/ɪ/, /ɛ/, /æ/, /ʊ/, /ʌ/) cannot occur in stressed open syllables (e.g. */bɪ/). Short vowels are more lax than long vowels and therefore, the terms tense and lax have been attributed to long and short vowels as (Plag, Ingo, Braun, Lappe & Schramm, 2007). The English vowels /iː/, /ɑː/, /ɜː/, /uː/, /ɔː/ can be considered tense, whilst all others are lax.

The same concept of vowel length contrast does not apply to Persian. Vowel length is more context dependent, for example vowels in CVCC context compared to CVC and CV syllables (Gharavian, Sheikhzadeh, & Ahadi, 2000, p. 189). In Persian, /ɑ/, /e/, and /o/ are considered as lax vowels and /æ/, /i/, and /u/ as tense vowels (Samareh 2000, p. 85; Windfuhr, 1979, p. 526). Hall (2007, p. 37) explains that variations in vowel length do not affect the meaning of words.

Apart from length contrast, English vowel pairs are phonemically contrasted in terms of vowel quality. Deterding (1997; 2003) shows that the vowel pairs /ɪ/-/iː/, /e/-/æ/, /ʌ/-/ɑː/, /ʊ/-/uː/, and /ɒ/-/ɔː/ are contrasted in Southern British English. For example, based on the first two formant frequency measurements (see section 2.3), /iː/ is realised higher and more front than /ɪ/ as is /e/ compared to /æ/. However, the degree of contrast is not consistent across the pairs. For instance, there appears less contrast between /ʊ/ and /uː/, and /ʌ/ and /ɑː/ in Southern British English (SBrE), based on their placement in the vowel quadrilateral (Deterding, 1997; 2003). In English, there is in fact interplay of both quality contrast and vowel length contrast (Cruttenden, 1994) and thus any study looking at vowel contrast should take both parameters into account. Cruttenden (1994, p. 92) explains that only in the case of /ə/-/ɜː/ can it be said that there is an opposition solely of length as "/ə/ occurs only in unaccented syllables, whereas /ɜː/ can occur in syllables carrying primary or secondary stress" (e.g. in *bird*).

2.2.2 Diphthongs

A diphthong is one continuous gliding vowel sound within one syllable and according to Jones (1990), it is produced when a sound is made by gliding from one vowel position to another. This nature of diphthongs is reflected in their phonetic transcriptions which has two symbols. Ladefoged and Maddieson (1995) also state that diphthongs are considered to be vowels that have two separate targets. Thus, a diphthong has two different target positions from onset to offset of the vowel and the changing transition between them occur quickly.

RP has eight diphthongs: /ɪə/, /ʊə/, /eə/, /eɪ/, /ɑɪ/, /ɔɪ/, /əʊ/, /ɑʊ/ (Gimson, 1970, p. 139-140); while American English is reported to have five diphthongs /ɑɪ/, /oʊ/, /ɑʊ/, /eɪ/, and /oɪ/ (Ladefoged, 2001b, p. 31-70). Some dialects of American English produce the vowels /e/ and /o/ as diphthongs, namely /eɪ/ and /ou/ such as in the words *fiancé* and *course*. RP has five closing diphthongs that end in /ɪ/ or /ʊ/, and three centring diphthongs ending in /ə/ as shown in Figure 2.3 and 2.4.

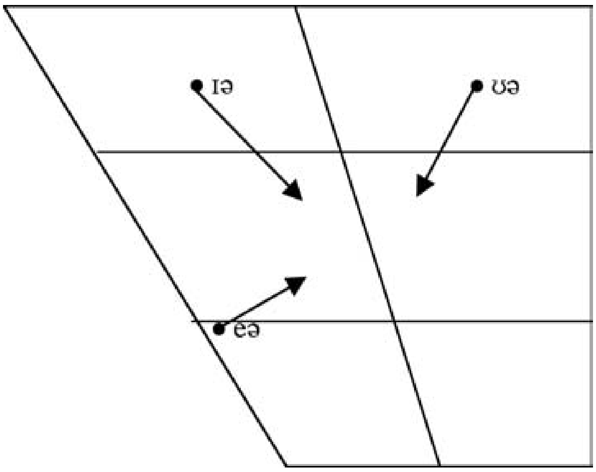

Figure 2.3: Centring Diphthongs
(Source: Roach, 2000, p. ix)

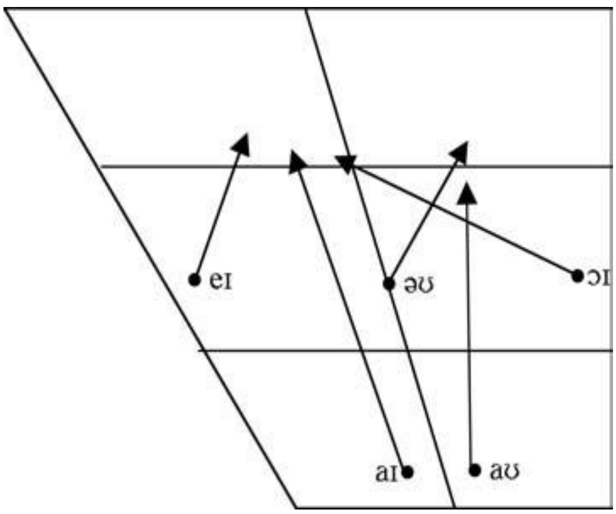

Figure 2.4: Closing Diphthongs
(Source: Roach, 2000, p. viii)

There are three diphthongs in modern British English which are not found in American English; /ɪə/, /eə/, and /ʊə/ (Gut, 2009, p. 62). In American English, Cruttenden (1994, p. 132-4) notes that /ɪə/ is realised as /ɪ/ or /iː/ + /r/ (for words spelt with a final r, e.g. *hear*); /eə/ tends to be realised as the long monophthong /eɪ/ or /e/ + /r/ (e.g. pear), and /ʊə/ is regularly realised as /uː/ or /ʊ/ː/ + /r/ (for /r/ before a consonant or a pause). Among the eight diphthongs, /ʊə/ is being replaced by the long monophthong /ɔː/, for example the words *poor* and *sure* are pronounced as /pɔː/ and /ʃɔː/ in Standard British

English (Jones et al., 2003, p. 420; Wells, 2000, p. 593). Roach, Hartman, and Setter (2006) also note that the long monophthong /ɔː/ is used instead of /ʊə/ in British English. However, words with /uː/ used with a suffixal /ə/ such as in *doer*, *fewer*, *newer*, *viewer* is never pronounced with the vowel /ɔː/ (Cruttenden, 1994). Further, Ladefoged (2001b, p. 31) finds that /ʊə/ is also replaced by /ɔː/ in American English. However, /ʊə/ is still listed as one of the vowels of British English in most textbooks, and thus, tends to be taught in the classroom, especially in EFL situations. In contrast to English, Persian only has three diphthongs; *ey*, *ow*, and *ay*. These three diphthongs are classified and pronounced similar to /eɪ/ in the English words *bait*, /əʊ/ in the English word *go* and /ɑɪ/ in the English word *sky* (Bashiri, 1991, p. 9). Thus, Persian and English have these three diphthongs in common: /eɪ/, /əu/, and /aɪ/. Keshavarz (2001, p. 9).

2.3 Formant Frequency Model

Formants are defined by Fant (1960) as the spectral peaks of the sound spectrum. Thus, formants are explained as constricted bands within the acoustic scope where energy is focused while the speech sounds are being produced; the rate of occurrences of formants is measured by the tenseness and resonances of vocal tract cavities (nasal, oral, pharyngeal). Vocal tract shape and size play role in changing formant frequency, therefore, changes produced by lip shape and tongue movement influence the formant frequency (Hayward, 2000; Watt & Tillotson, 2001). Formants are visible on a wide-band spectrogram as broad, dark bands extending across the duration of a vowel, and the patterns of the formants, particularly the first two formants, are utilised to analyse vowels (Fry, 1979; Hayward, 2000). These two formants reveal the vowel features regarding open/close and front/back dimensions; although Kent and Read (2002) caution that there will be exceptions.

The first formant frequency (F1) is related to vowel height as the higher vowel will be lowered in the frequency. The second formant frequency (F2) is related to tongue retraction so that those vowels with higher F2 will be the more advanced vowels (Fry, 1979; Kent and Read, 2002). Therefore, a high, front vowel like /i/ is likely to contain a high F2 and a low F1 while an open back vowel like /ɑː/ is likely to have a higher F1 frequency and a low F2 frequency (Ladefoged, 2001b; Watt & Tilloston, 2001). In the spectrogram in Figure 2.5, the blue arrow indicates the first formant for /i/ and the second formant is shown by the green arrow. For /ɜ/, the red arrow points to second formant which occurs at a higher frequency compared to /i/. For /ɑ/, the orange arrows point to the first and second formants indicates that they are fairly fused; therefore making it difficult to decide where exactly they lie. As Kent and Read (2002) explain, "[b]ack vowels have a low F2 and typically a small F2-F1 difference, whereas front vowels have a relatively higher F2 frequency and a large F2-F1 difference" (p. 113).

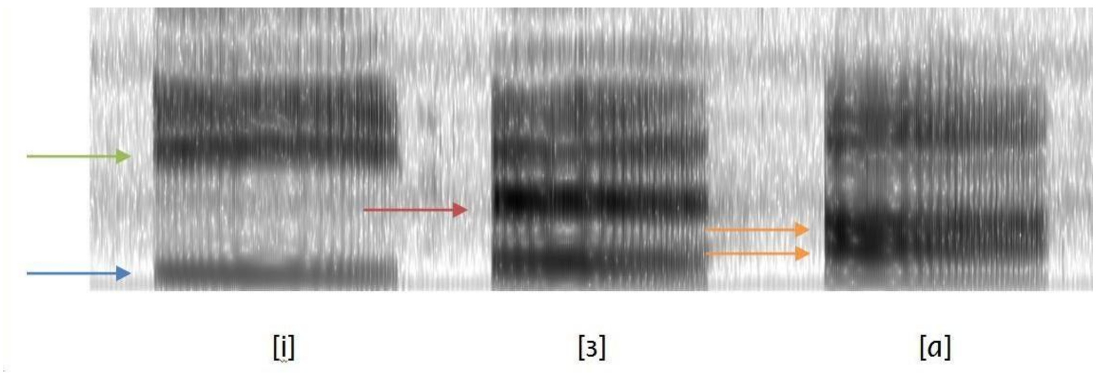

Figure 2.5: F1 and F2 for /i/, /ɜ/, /a/
(Source: Deterding, 2008, from http://Knol.google.com)

To determine the first and second formants and to indicate the place of them, formant tracks are overlaid over the spectrograms. In the spectrogram in Figure 2.6, the formant tracks are shown in red.

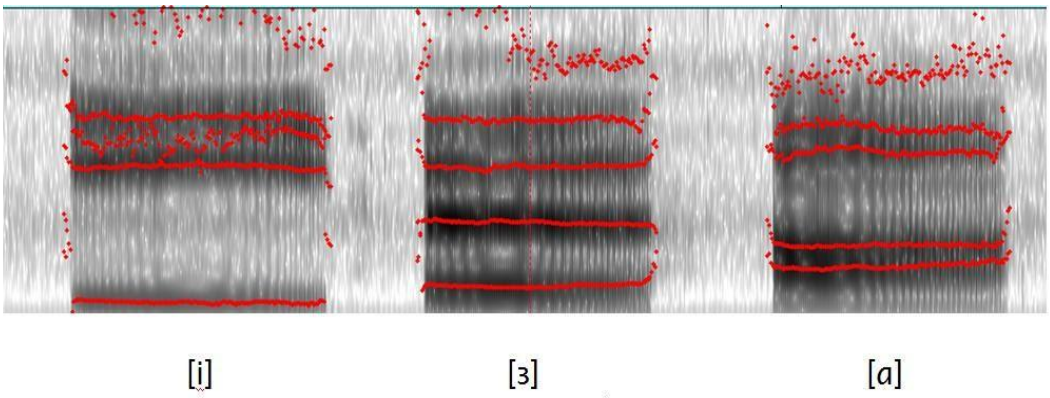

Figure 2.6: Formant Tracks
(Source: Deterding, 2008, from http://Knol.google.com)

As the first two formants of a vowel are considered important for the perception of vowels, most studies on vowels tend to focus on the first two formants (e.g. Deterding, 2003; Fleming & Johnson, 2007; Hawkins & Midgley, 2005; Ladefoged, 2001b), although some studies like Yan & Vaseghi (2003) also considered the third and fourth formants. Yan and Vaseghi (2003) examined the formant space of three major English accents namely British, American and Australian. They measured the first to fourth formants and found that a comparison between the averages of these formants showed that "formant shifting exerts an important impact on accent". With the third formant providing more information about differences in accents (p. I- 715).

Childers and Wu (1991) go even further and measure the average of the first, second, third and fourth formants because they were of the opinion that "the accuracy of formant tracking using the conventional frame-based LPC (Linear Predictive Coding) analysis is affected by following factors (page. I-713-714):

1) Influence of pitch on the first formant.

2) Formant movements resulting in the merging of the trajectories of adjacent formants.

3) Rapid formant variation that may occur in consonant vowel transitions or diphthongs.

4) Source-vocal tract interaction (ignored in LPC analysis)

5) Effects of lips radiation and internal loss on formant bandwidth and frequency.

(ibid., p. 1841-1856)

However, in instrumental phonetics, most researchers have adopted the formant frequency model in which vowel quality is typically analyzed through the first two formants patterns. According to Watt and Tillotson, the formant frequency model is used to analyze the vowels because: "the formants contain most energy during sonorant sounds such as vowels" (2001, p. 275). The rate of recurrences of F1 and F2 in relation to one another facilitates and determines the required signs for the identification of individual vowel qualities. Furthermore, F1 and F2 frequencies correlate with tongue position; in other words, one can say that the rise of F1 frequency is consistent with lowering of the tongue and opening of the jaw, however a rise in F2 frequency originates from fronting of the tongue body (e.g. Ladefoged and Harshman, 1979).

Pillai, Mohd. Don, Knowles & Tang (2010) in their study on Malaysian English admit that the formant frequency model has some limitations but submit that this model is still the most used one in instrumental analysis of vowels. The average F1 and F2 measurements are then plotted on a F1 versus F2 plot. Watt and Tillotson (2001, p. 276), opine that "the formant plots can provide an approximate representation of the relative qualities of individual vowels for single speakers". The F1 /F2 plot can give a better representation of the "traditional vowel quadrilateral" than the F2−F1/F1 plot (Watt and Tillotson, 2001, p. 213). This method allows the researcher to overlay one speaker’s vowel sample on top of the other. However, according to Hayward (2000) there is an issue regarding the positioning of the back vowels because of lip rounding which has an effect of causing a dip in the F2 (see Fry, 1979). This impact is intensified in English, as four of the back vowels have some degree of lip rounding. Thus, measurements of the back vowels need to be evaluated more carefully.

It should be noted that the average F1and F2 values for the monophthongs are usually converted into Bark scale (Zwicker & Terhardt, 1980). Many studies related to instrumental analysis of vowels employ the Bark scale because "it is thought to be a good approximation of the actual frequency analysis performed by the ear" (Kent and Read 2002, p. 115). As Deterding (1997, p. 50) expounds, "[t]he main effect of representing the front/back dimension in terms of F2 - F1 would be to normalize for speaker differences, particularly male-female differences in formant frequencies".

Further, Salbrina (2009, p. 49) explains that she uses the Bark scale in order:

> [t]o obtain a vowel plot that is “psychologically real” (Hayward 2000, p. 152) and in order to achieve this, the axes of the plots should correspond to the perceived pitch of the sound rather than its physical frequency because the relationship between the frequency and the perceived vowel sound is not linear.

2.4 Measuring Diphthongs

Although Ladefoged (1993, p. 196) mentions that there is a strong relation between "vowel openness" and the first formant (F1), it is difficult to determine the best way for describing diphthongs acoustically since English diphthongs are characterized by a changing vowel quality. One of the ways for the measurement of diphthongs is the one used by Deterding, (2000), where the Rate of Change (ROC) of the first formant (F1) is derived as an approximate estimate for the quality of diphthongs. The use of ROC for the measurement of formants is recommended by Gay (1968) and Kent and Read, (1992). Since a diphthong is composed of an initial steady state (onset), a final steady state (offset), and a transition between them, ROC is measured by finding the difference in the formant values at the beginning and at the end of a target diphthong (e.g. onset and offset of the diphthong) and then dividing the resulting value by the duration of the diphthong (see Deterding, 2000; Salbrina, 2006).

It is necessary to mention that other researchers also have proposed more complex acoustic ways to measurement and description of diphthongs. Ren (1986), for example measures the F2 trajectory at various points in the diphthong, while Clemont (1993) suggests taking the third formant into account. Deterding's (2000, p. 96) suggested algorithm is used by Deterding, Wong, and Kirkpatrick (2008), Pillai (forthcoming), Salbrina (2006) and Tan and Low (2010) because while the simplistic use of ROC of F1 is enough to give an indication of the extent to which there is diphthongal movement. Deterding (2000) employed this method of measurement in his study on diphthongs in Singapore English (SgE) which focused on /eɪ/ and /əʊ/. Data derived from ten female Malay Singaporean speakers, ten female Chinese Singaporean speakers, and five female British lecturers were analysed and he found that of the ROC for the British speakers was higher than the Singaporean speakers indicating that the former were producing the vowels with more diphthongal movement than the latter. However, Lee and Lim (2002) found that only measuring the ROC for F1 was not sufficient for centring diphthongs like /eə/ since height is not a factor here. Thus, Lee and Lim (2002) also measured the ROC for F2 to reflect the front/back movement. Further, to provide an illustration of the diphthongal movement, Deterding, Wong, and Kirkpatrick (2008) plotted the F1 and F2 from the onset to offset of the diphthongs /eɪ/ and /əʊ/. Such plotting shows the trajectory of the diphthongs, providing a picture of the direction and the extent of the movement.

2.5 Instrumental Studies on English Vowels

Hillenbrand, Getty, Clark and Wheeler (1995) conducted an instrumental study on American English elicited from male, female speakers and children. To record the data, the subjects were asked to read 12 vowels /i/, /ɪ/, /ɛ/, /æ/, /ɑ/, /ɔ/, /ʊ/, /u/, /ʌ/, /ɜ/, /e/, /o/ and four diphthongs in hVd context. The first to fourth formants were measured based on the LPC spectra. In comparison with other studies conducted on American vowels (e.g. Di Benedetto, 1989; Peterson & Barney, 1952), Hillenbrand et al., (1995) found higher F2 values and lower F1 values for /æ/ than /ɛ/. A major difference between their findings and Deterding's finding showed that there is a high degree of overlap in F1 and F2 space for /æ/ and /ɛ/ for American vowels while in British vowels this kind of overlap is not considered. Figure 2.7 shows the F1 and F2 values for American vowels.

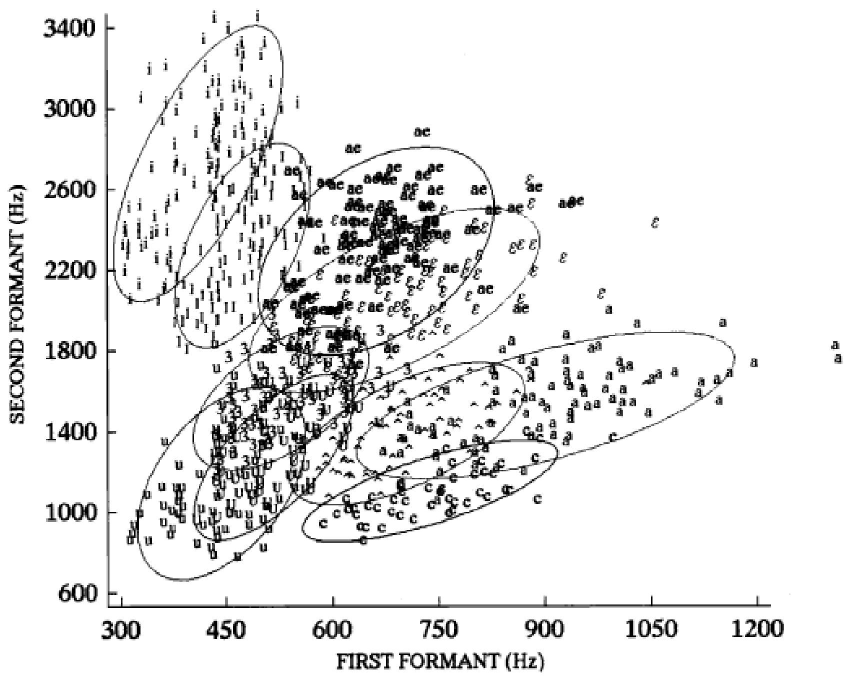

Figure 2.7: Values of F1 and F2 for 46 Men, 48 Women, and 46 Children for 10 American English Vowels
(Source: Acoustic Characteristics of American English Vowels, Hillenbrand et al., 1995, p. 3104)

Cox (2006) on an acoustic study about characteristics of Australian English (AusE) vowels also used an hVd context. She found that there are no differences or contrast between the /ɪə/-/ɪ/, /eː/-/e/, and /ɑ/-/ɑː/ vowel pairs. Her data were derived from 60 Australian male and female teenagers who produced vowels. She also analysed the diphthongs /ɑe/, /æɪ/, /oɪ/, /æɔ/, /əu/, because "there is evidence in the literature that these vowels are becoming progressively more monophthongized in South Australian English (SAusE)" (Bradley 1989; Horvath 1985, cited in Cox, 2006, p. 153). The results from the formant trajectory diagrams (p. 159-161) confirmed this evidence that these diphthongs for SAusE are produced more monophthongal than diphthongal. The findings for Australian monophthongs showed similar results between both male and female subjects. For both monophthongs and diphthongs, Cox measured F1-F3 values. The results for diphthongs showed that there were slight differences in producing diphthongs among male and female speakers with female subjects producing diphthong

/ɪə/ with different first target vowels (see Cox, 2006, p. 156-161). However, the findings also demonstrated that monophthongs for males are shorter than those for females. Similar to findings on a study on American vowels (Hillenbrand et al., 1995), females appear to be producing longer monophthongs and also displayed a longer transition for diphthongs.

Using the same word context, a similar study was also conducted by Ferragne& Pellegrino (2007) on the monophthongs of East Anglia, England. This study was carried out to compare the monophthongs of East Anglia speech with standard British English. The data were derived from formant frequency measurements on 11 vowels also in hVd contexts. The data analyzed came from the accents of the British Isles (ABI) corpus. The subjects for this study were 19 speakers; ten female and nine male, who produced five series of nineteen vowels in hVd contexts. The subjects were asked to read "*heed, hid, head, had, hard, hod, hoard, hood, Hudd, heard, who'd"* in random order. Using words in hVd contexts are common in different phonetic studies since the preceding fricative and final stop consonant aids the identification of the target vowel on the spectrogram and have minimal co-articulatory effect on the vowel (e.g. Cox, 2006; Cox & Palethorpe, 2005; Hillenbrand et al., 1995). However, one of the limitations of hVd words was that some phonological contrasts might be missed. The findings were compared with previous published studies on standard British English. Ferragne found that back and central vowels in East Anglia monophthongs were fronted and more open than standard British English but that the front vowels were similar.

Other studies used different word contexts to host target vowels. For example, Hubais and Pillai (2010) who conducted an instrumental analysis on English monophthong vowels produced by Omani speakers used a word list in a CVC, where C was a voiced stop consonant F1 and F2 were measured based on LPC and average values for both formants and converted into Bark scale to examine quality contrast between vowel pairs. Vowel length was also measured and compared. The results showed that the vowels /ɪ/ and /e/ overlapped, and that the vowel /e/ was pronounced as /ɪ/ (Hubais & Pillai, 2010, p. 9). The authors found that the vowels /iː/, /ɪ/, /æ/, /ɑ/, /ʌ/, /ɜː/, /ʊ/, /ɔː/ were produced more fronted than British English and found a lack of length contrast between vowel pairs /ɪ/-/iː/, /ʊ/-/uː/ and /ɒ/-/ɔː/. In general, the results indicate that Omani speakers produced less peripheral vowels compared to British English.

In terms of vowel quality contrast, the findings in an instrumental analysis on Malaysian English by Pillai, Mohd. Don, Knowles and Tang (2010) showed a lack of quality contrast between the vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʌ/-/ɑː/ (p. 165). The data comprised recordings by 47 female Malaysian undergraduates**.** Like Hubais and Pillai, the target vowels embedded in /CVC/ context. The lack of quality contrast between vowel pairs was similar to Singapore English (Deterding, 2003) and Brunei English (Salbrina, 2006) except for /ɒ/ and /ɔː/.

Apart from word lists, other speaking contexts used are texts and spontaneous speech. Salbrina (2006), for example, obtained her data from ten female Brunei speakers and seven British female speakers who were asked to read *The North Wind and the Sun* (NWS) text. Salbrina (2006, p. 249) explains that "[t]his passage was chosen because this passage has been traditionally used by the International Phonetic Association [IPA, 1999, p. 39] in phonetic analyses of language". Apart from analyzing monophthong vowels extracted from the text, Salbrina also examine the diphthong /eɪ/ by measuring the ROC. One of the problems of using the NWS text was that she had to ignore words in which the vowels were followed by approximants /j/, /w/, and /r/ and preceded by the velar nasal /ŋ/ and dark /l/ to avoid the co-articulatory effects of other consonant sounds. The result showed that Brunei /uː/ and /ʊ/ are similar to Standard British vowels and more fronted than Singapore ones while Brunei /ɔː/ is more open compared to British and less back. The ROC value for the Bruneian diphthong /eɪ/ indicated that the average ROC for Brunei English was less than the British /eɪ/ in her study, meaning that the Brunei vowel has less diphthongal movement. Salbrina's findings suggest that Brunei English speakers produced this diphthong as a long monophthong /e/ similar to Singaporean speakers in Deterding's (2000) study.

A comparison between Malaysian and Singapore English monophthongs (Tan & Low, 2010) showed that Malaysian English vowels were more compact in the vowel space than Singapore English vowels. Similar results were found in Pillai, Mohd. Don, Knowles and Tang (2010). In Tan and Low (2010), data were collected from five male and five female speakers from each variety (total of 10 males and 10 females), all undergraduate students with similar background. The subjects were asked to read *The Boy who Cried Wolf* (Wolf) texts and also asked to read citation forms where, the target vowels were embedded in a CVC word which were placed in a carrier frame *Please say CVC again*. The reason for using the *Wolf* context instead of NWS (the standard one) explained by Tan and Law (2010, p. 168) as "Deterding (2006) demonstrated that the Wolf passage works well for the description as well as the measurements of vowels and consonants in English". In terms of vowel length, the results from citation forms showed that Singapore English speakers distinguished vowel length between vowel pairs while

Malaysian English speakers did not differentiate the length for the vowel pair /ɒ/-/ɔː/. The values taken from vowel plots implied that there was no contrast between vowel pairs /ɪ/-/iː/ and /ʊ/-/uː/ for both varieties. The vowel /ɔː/ was produced less back and more open among Malaysian English speakers as the values of F1 and F2 showed higher values compared to Singapore English speakers.

Another study on Hong Kong English (Deterding, Wong & Kirkpatrick, 2008) also reported similar findings to Singapore English, Malaysian English and Brunei English. The Hong Kong English vowel pairs /ɪ/-/iː/, /e/-/æ/, and /ɒ/-/ɔː/, were closer than in British English. The Hong Kong data were taken from three female and three male educated speakers. A major difference with the other Asian varieties was the position of /uː/ which was much further back position than in Singapore English but more fronted in Hong Kong English. Deterding, Wong and Kirkpatrick (2008) assumed that the reason for the fronted position of /uː/ is not due to the influence of the L1 of speakers (Cantonese) on, but rather the influence of British English, which also displays similar tendency to front this vowel. A study of formant frequencies (F1 and F2) on RP monophthongs by Hawkins and Midgley (2005) confirm this phenomenon and this has also been found in New Zealand English (Bauer and Warren, 2004).

Deterding's Singapore English data and the Hong Kong English study were derived from an informal speech context. The use of informal speech results in a more naturalistic result because it is not controlled by the careful production. However, not all the target sounds will be obtained and even if they do there is no control over the frequency of occurrences. Choosing only the vowels in stressed syllables and content words limit the study (e.g. Harrington 2006; Jacobi, Pols & Stroop 2006). This is why many studies combine spontaneous speech with a more controlled speaking context (e.g. Deterding, 1997). Thus, the informal speech cannot be used alone or directly compared to the frequency of occurrences of English vowels (e.g. Cruttenden 1994; Knowles 1987) because as Knowles (1987, p. 223) points out, "frequency distributions of phonemes will vary according to the different varieties of English and the types of text from which the frequencies are obtained".

2.6 Studies on Vowels Produced by EFL Learners

Attempts have been made to analyse English vowels produced by EFL learners Among them is an acoustic analysis on Japanese- accented English by Tsukada (1999) who compared the vowels produced by Japanese speakers with native Australian English speakers in terms of vowel quality. Tsukada found that the Japanese speakers had higher values for F1 and F2 than the Australian ones. The subjects comprised of 12 (six male and six female) Australian English speakers and 25 (11 male and 14 female) educated Japanese speakers and was divided into two parts. The first part comprised eight English monophthongs /iː/, /ɪ/, /e/, /æ/, /ɒ/, /ʊ/, /ʌ/, /uː/ in CVd and CVt word contexts. The second part comprised five Japanese vowels in CVdo and CVt o word contexts. The subjects were asked to read words list in a random order while being recorded. The F1 and F2 values also measured (see Tsukada, 1999, p. 377). Japanese male and female subjects produced English vowels similar to each other as findings for both of them showed a lack of contrast among vowel pairs /ɪ/-/iː/, /ʌ/-/ɑː/ and /ʊ/-/uː/, a trend which is also common in second language varieties in Southeast Asia. However, for both males and females there was a merger of /ɑ/ and /ʌ/ vowels. Compared to English, Japanese has much simpler vowel system as there are only five vowels /i/, /e/, /ɑ/, /o/, and /u/ in Japanese that contrast in vowel length. Tsukada also found that in terms of vowel duration, Australian speakers produced longer vowels than the Japanese ones. Tsukada concluded that Japanese subjects were strongly under the influence of L1 as findings on durational differences between the vowels showed the larger ratio (1 to 1.59 ms) in Australian English and the least ratio (1 to 1.06 ms) for Japanese English.

Another study compared English vowels produced by Swiss EFL speakers with British and American English native speakers (Leemann, 2007) vowels. Leemann (2007) performed an acoustic analysis on 10 English vowels: /iː/, /ɪ/, /uː/, /ʊ/, /ɑː/, /ʌ/, /ɒ/, /æ/, /e/, /ɜː/. The formant values were taken from Received Pronunciation (RP) vowels (Deterding, 1997) and General American English (GA) vowels (Hillenbrand et al., 1995). The recordings of the EFL subjects comprised of 20 secondary school students reading word lists in an hVd context. The results showed that the Swiss English vowels /ʊ/, /uː/, /ɑː/, and /ɪ/ were produced farther from the target vowel than the Swiss English vowels /iː/, /ɛ/, /ɜ/, /æ/, /ʌ/ (see p. 9). This result was because of higher F1 values taken for the Swiss English vowels, /ɑː/, /ʊ/, /ɪ/, /æ/, /ɜ/, /ʌ/, /iː/, /uː/, and a lower F1 values for vowels /ɛ/ and /ɔ/ compared to RP vowels. While, the F2 values for Swiss English vowels showed further backness for /iː/, /ɜ/, /ʌ/, /ɑː/, /uː/ and further fronted vowels /ɪ/, /ɛ/, /æ/, /ɔ/ compared to RP. A comparison between GA and Swiss English showed that /ɛ/, /ɜ/ and /ʌ/ were similar to each other while, the other vowels in terms of vowel space distance showed greater values of F1 and/or F2. The differences between F1 and F2 showed that Swiss English vowels were produced further back, except for the vowel /ɜ/ that was produced more fronted compared to GA (p. 10). There was a lack of contrast between /ʊ/-/uː/ and /e/-/æ/ vowel pairs. The vowel /e/, /ɜː/ and /ʌ/ were similar in Swiss

English and GA. In terms of vowel height, seven out of ten Swiss English vowels /iː/, /ɪ/, /ɑː/, /e/, /ɔː/, /uː/, /ʊ/ showed more similarity to RP pronunciation, while these three vowels /ɜː/, /æ/, /ʌ/ were pronounced similar to GA (Leemann, 2007, p. 9-11).

2.7 Summary

This chapter has shown that there are differences and similarities between English vowels in different varieties of English, which inadvertently influence the production of English vowels by native and non-native speakers despite the fact that there is still no satisfactory explanation as to whether similar categories will lead to more or less difficulty in learning a particular sound. The formant frequency model is a much used method for measuring and analysing vowel duration.

In terms of Persian speakers' production of English, there is, at present, a scarcity in instrumental studies on the realization of vowels by Persian speakers. Thus, the current study will attempt to fill this research gap. In particular, it will contribute to the understanding of how Persian speakers produce English monophthongs and diphthongs. In relation to this, the following chapter describes the methodology used in this study.