CHAPTER 2

REVIEW OF RELEVANT LITERATURE

2.0 Introduction

In this chapter, relevant literature will be reviewed, beginning with the presentation of the respective characteristics of vowel systems among the languages of Arabic, English and Malay before moving on to a description of Arabic and Malay-accented English, followed by perusing theories corresponding to some of the important factors that influence foreign/second language speech perception. Previous empirical studies will also be discussed. Among all these issues of reviewed researches, the peculiar characteristics of the English vowel monophthongs realized by Omani speakers are discussed in detail, which will be applied in the next chapter of data analysis in order to provide indications and solid background knowledge for interpreting the results of the perception tests of Malay subjects of English towards Omani production together with the impingement of the actual realization of Malaysia English.

2.1 Vowels in Arabic and English

Vowels in general refer to vocoid sounds produced without closure or narrowing typical for the consonantal sounds (Gimson, 1994). “Linguists are often faced with the problem of comparing vowels” (Ladefoged, 1967, p. 52) since the lip and tongue gestures for producing them vary in different languages; and the number and quality of vowel systems of different languages are not identical either.

There are 12 monophthongs - /ɪ/, / iː/ /e/, /æ/, /ɜː/, /ʌ/, /ɑː/, /ɒ/, /ɔː/, / ʊ/, /uː/, /ə/ - and eight diphthongs - /aʊ/, /əʊ/, /eɪ/, /aɪ/, /ɔɪ/, /ɪə/, /eə/, /ʊə/ in the English language (Roach, 2000), whilst Standard Modern Arabic (SMA) possesses a comparatively small inventory of vowels (see chapter 1.3). Table 2.1 shows the eight Arabic vowel phonemes, where it can be seen that each of the diphthongs is composed with a vowel that is immediately followed by a semi-vowel, so that there is a glide in the syllable.

Table 2.1

The Vowel Phonemes of Arabic

No.	Vowel	Definition	Example	
			Phonetic Transcription	**Meaning**
1	/i/	Front high unrounded long	/sʸit/	reputation
2	/ɪ/	Front high unrounded short	/sɪt̪/	six
3	/u/	Back high rounded long	/sʸum/	fast (v.)
4	/ʊ/	Back high rounded short	/sʸʊm/	deaf
5	/æ/	Front low unrounded long	/mæl/	money
6	/a/	Central low unrounded short	/mal/	he got bored.
7	/ay/	*See the paragraph below*	/kayl/	how
8	/aw/	*See the paragraph below*	/lawn/	color

(Sauce: Khaled Huthaily, 2003, p. 29)

Besides the shared use of SMA, people in Arab countries manipulate their own dialects. Alghmdi (1981) looked at the Saudi, Sudanese and Egyptian varieties of Arabic, and found an explicit diversity of quality change in the vowels. Even though there were

quality differences in the vowels of the three dialects, he discovered the coincidence that all of them maintained length contrast between vowel pairs (cf. Alghamdi, 1998, p. 21). Flege (1979) also examined the duration of vowels in Arabic with English and found that the long vowel /aa/ in Arabic is closer in duration to English /æ/ than is short Arabic /a/. He asserted that in a pre-dental stop environment the duration of Arabic was 177 msec, short Arabic, 98 msec, whilst English averaged 187 msec when produced in a comparable phonetic context by Americans. Similar rules are adopted in English as well, that apart from the quality difference, the length between vowels also plays an important role in distinguishing different vowel pairs (Ladefoged, 2006). For example, in order to distinguish the pair /ɪ/ and /i/, both the differences regarding length and quality have to be taken into consideration.

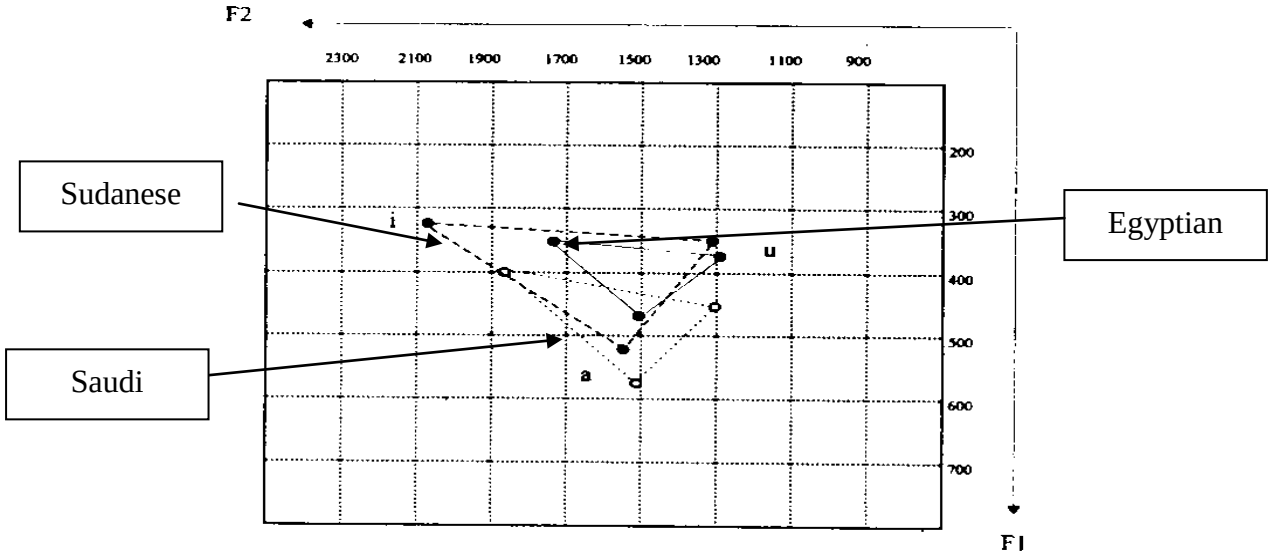


Figure 2.1: Formant Chart of Saudi, Sudanese and Egyptian Vowels

(Source: Alghambi, 1998, p. 21)

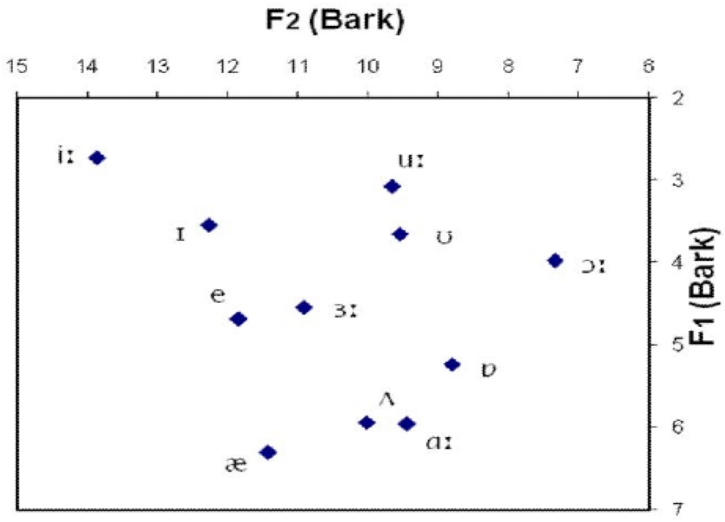


Figure 2.2: Plotting Vowel Quality of English Monophthongs

(Source: Deterding, 1997, from
http://knol.google.com/k/measuring-and-plotting-vowels#)

Subscription to the discoveries of Alghmdi (1981) mentioned above, previous studies on varieties of English indicated differences in the realization of vowels. In a study of Brunei English, Salbrina (2006) reported a lack of contrast in the vowel pairs /i:/ and /ɪ/, /ɔː/ and /ɒ/, and /e/ and /æ/. The speech of English of Malaysia and Singapore has been analyzed as a continuum ranging from the acrolect to basilect (Platt & Webber, 1980), where it is asserted that the RP pair /e, æ/ was often conflated to /ɛ/ (Brown, 1997).

2.2 The Production of English Sounds by Native Arabic Speakers

Generally the English sounds produced by native Arab speakers bears similar characteristics, Flege and Port (1981, p. 125) postulating that “perhaps the most important and obvious aspect of foreign-accented speech is sound substitutions”. Based on this they found native speakers of English produced longer VOT but shorter stop closure intervals in English than the Arabic L2 speakers of English, which was constant

for the findings in the Arabic experiments, where the sum of the VOT and stop closure intervals of the same group of Saudi speakers were tested. It indicated that there was L1 interference in the native Arabic speakers' production of English. Mitleb (1981) also reported a transfer of Arabic long-short vowel duration patterns to English tense and lax vowel pairs of Jordanian speakers. At the mean time, it is proposed that "the Omani speakers are influenced by Arabic vowels in their production of similar English vowels".

The English used in the Arabian Peninsula bears its own peculiar characteristics according to investigations by researchers in comparison with the standard varieties of the English language. Wells (1982: 75) proposed, "accents clearly vary in the details of segment duration", while Hubais (2009) in his study of Omani English speakers drew conclusions from examining the three vowel pairs /iː/and /ɪ/, /uː/ and /ʊ/, and /ɔː/ and /ɒ/ that "Arabic speakers tend to maintain, if not exaggerate, length contrast even if vowel quality is not strongly contrasted". This is consistent with previous studies of Alghamdi (1990, cited in Alghamdi, 1998) as Arabic was found to have a short to long vowel ratio of 0.51. Munro (1993) distinguished between English vowels produced by Arabic speakers and native American speakers of English, and discovered that on average, the realizations of all the vowels by the former were consistently shorter than those produced by the latter. Subsequent to this, he provided explanation for this phenomenon, which could be due to first language interference. In this research, he also declared that there was a tendency for Arabic speakers of English to produce the English tense and lax vowel pairs similarly with the vowel pairs in Arabic, which

resulted in a significant difference in the durational ratios of their production and that of native English speakers. Findings of another research also showed in terms of the duration of English vowels produced, “the effect of stop voicing on vowel duration is much smaller in the Saudis’ than in the Americans’ English” (Flege and Port, 1981, p. 137), which contradicted the reported findings for English that vowels were always produced longer before voiced than voiceless stops by native speakers.

Moreover, concerning vowel quality, Hubais (2009) investigated the English monophthong vowels of Omani speakers, which agreed with the findings of previous studies. For example, /e/ is merged into /ɪ/, a phenomenon which was also reported by Munro (1993). Thus, when the word *pet* was played, the word *bit* and *pit* was perceived; whilst *pit* and *pet* were perceived as *bit*, according to the findings of a perceptual test of Hubais (2009), which also indicated a tendency of realizing the voiceless stop as the correspondent voiced one. A similar phenomenon is also evident in Flege & Port (1981) and Rasmussen (2007). In the same study there is also a comparison of formant frequencies of English monophthong vowels produced by Omani speakers and British speakers. It is reported that those produced by Omani subjects “occupy a slightly more compact vowel space than British English”, as shown in Figure 2.3, where the average F1 values of “/iː/, /ɪ/, /ʊ/, /uː/ and /ɔː/ were produced lower in the vowel space by the Omani speakers, while /æ/ and /ʌ/ are produced higher than in British English”. A comparison of F2 values are plotted in Figure 2.4.

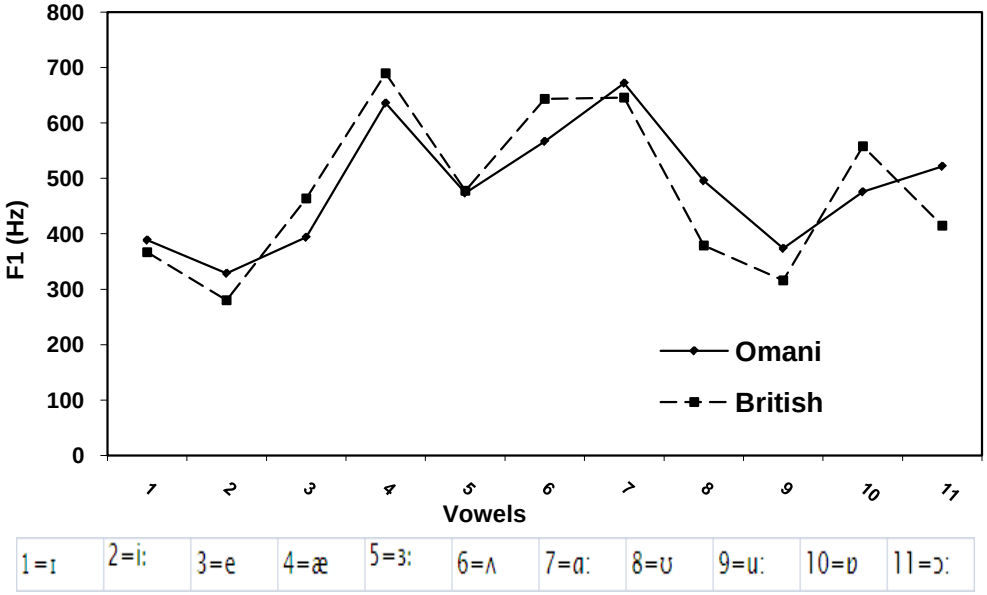


Figure 2.3: Comparison of F1 between Omani and British English Vowels

(Source: Ali Hubais, 2009, p. 59)

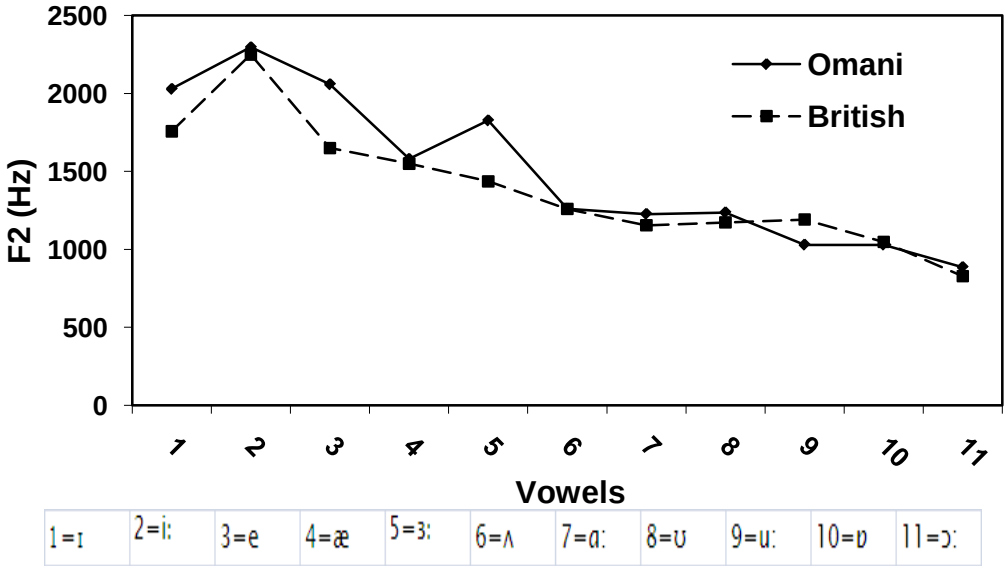


Figure 2.4: Comparison of F2 between Omani and British English Vowels

(Source: Ali Hubais, 2009, p. 60)

2.3 English in Malaysia and the Production of Vowels in Malaysian English

Malaysia is a multi-racial country; besides Malay peoples, other ethnic groups share its geographical location. Consequently, a variety of languages are manipulated and contribute to form a diversity of culture in the society. Together with the official national language, Bahasa Malaysia, which is used as the primary medium of instruction in education at all levels, coexist the Chinese language (Mandarin) and Tamil that share a vernacular status (Baskaran, 2005), though there are other languages in actual currency amongst the people of Malaysia. However, English is considered "as a strong second language" (Loga Baskaran, 2005, p. 15) and is considered crucial in "educational instructions, as well as… in the professions" (Asmah Omar, 1997, p. 12).

2.3.1 Background of Malaysian English

Since the British intervention in the Malay peninsula in the 18th century, the English language has gone through a process of "indigenization" and "institutionalization" (Moag, 1992) and developed its own variety, known as Malaysian English (Baskaran, 2005). Gaudart (1997, p. 47) reported that Malaysia English has many varieties, ranging from the "pidgins of hawkers" which is indigenized to the standard form of English spoken in myriad local accents used for official purposes, which is Standard Malaysian English (SME). Subscribing to previous assertions, parallel classifications have been made by Baskaran (1994, p. 27) as well that ME is perceived as a "continuum", with at least three distinguishable sub-varieties: the acrolectal, mesolectal and basilectal varieties, as shown in Table 2.2.

Table 2.2

The Malaysian English Continuum

	⟷		
	Acrolect	**Mesolect (Colloquial)**	**Basilect (Broken)**
S	**+ standard**	**- standard**	**Extremely simplified structures**
L	+ localized lexical items accepted in formal and informal use	+ localized lexical items, including those not used in more formal contexts	Pidgin-like
P	Can be ± marked ethnically ethnic accent and intonation	Usually, but not necessarily + marked	Usually + marked ethnic accent and intonation
E.g	• Newspaper reports • Formal letters and documents • Televison news • Official speeches	Informal spoken & written communication between colleagues, friends, family members etc.	Used by those with limited proficiency in English

(Source: Stefanie Pillai and Fauziah Kamaruddin, in "Varieties of English in Southeast Asia and Beyond", p. 39-54, 2006)

In the spoken of Malaysian English (ME), there may be first language (L1) influence in the pronunciation of speakers (Zuraidah Mohd Don, 1997, p. 35). As ME is greatly influenced by other vernacular languages, especially Bahasa Malaysia, it is associated with particular linguistic features, which will be discussed in the following section. Apart from influxes of localized lexis, there are also reported distinctions in its phonological features and intonation patterns along with differences in syntactic structures and pragmatic features (Platt, Weber and Ho, 1983; Anthonysamy, 1997; Nair-Venugopal, 2001). For example, "the dental fricatives /θ/ and /ð/ are often substituted by the corresponding alveolar stops /t/ and /d/ respectively" (Baskaran, 2005, p. 26) for there is a lack of dental fricative in Bahasa Malaysia. In terms of

pronunciation, the standard variety of ME is said to be similar to the Received Pronunciation of British English (RP) for the country's history of being colonized by the United Kingdom.

2.3.2 Production of Vowels in Malaysian English

As Zuraidah (1997, p. 35) asserted, "the more different the mother tongue is from the target language the more difficult it is to pronounce the language the way a native speaker does"; thus the phonological inventory of ME can be assumed to bear its distinctive features. For example, in Standard Malay (SM) there is a lack of "length differentiation among its vowels" (Zuraidah, 1997, p. 36), which is shown in Table 2.3; that is, it is reported that speakers of ME do not distinguish vowel length (Baskaran, 1994; Platt and Weber, 1980). Another commonly postulated feature of ME is that there is a tendency to produce monophthongs in place of RP diphthongs (Rajadurai, 2004; Wong, 1984), for instance, the diphthong /əʊ/ in *so* is realized similar to the vowels in *hawk* and *caught*, which is /ɒ/, but shorter (Pillai, Zuraidah, Knowles & Tang, 2010); /ei/ is said to be produced similar to the RP /ɛ/ or /ɛ:/ (Platt and Weber, 1980, p. 172-173).

Table 2.3

Vowel Phonemes of SM

	Closed Syllables		**Open Syllables**	
/i/	minta	pindah	nanti	pasti
/e/	tembuk	benteng	petah	elok
/a/	nampak	lapar	sapa	lalat
/o/	potong	gemuk	botol	sotong
/ə/	sembah	gempar	siapa	emak
/ai/			abai	capai
/aʊ/			pulau	saudara
/oi/			amboi	sepoi

(Sauce: Zuraidah, 1997, p.46)

Pillai, Zuraidah, Knowles & Tang (2010) conducted an instrumental analysis on English monophthong vowels produced by Malaysian speakers. In their study, there were 47 female speakers of different ethnic backgrounds recruited, whose productions of 517 tokens of CVC context with 11 target vowels embedded. A thorough description of the qualities of ME monophthongs were plotted as shown in Figure 2.5, which indicates a lack of contrast between /ɪ/, /iː/, /e/, /æ/ and /ʌ/, /ɑː/; whilst the vowel pair /ɒ/, /ɔː/ displays more contrast.

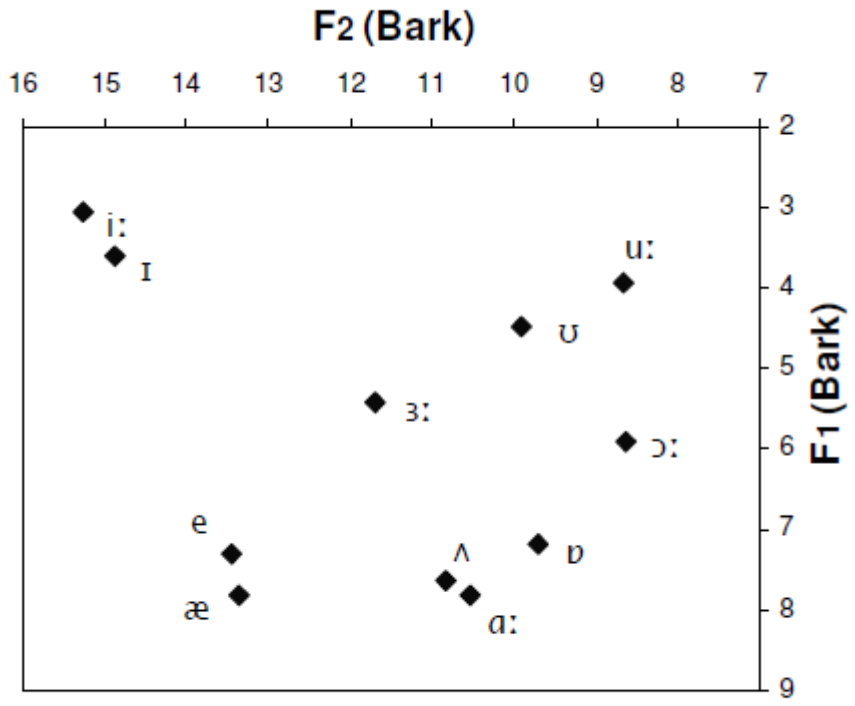
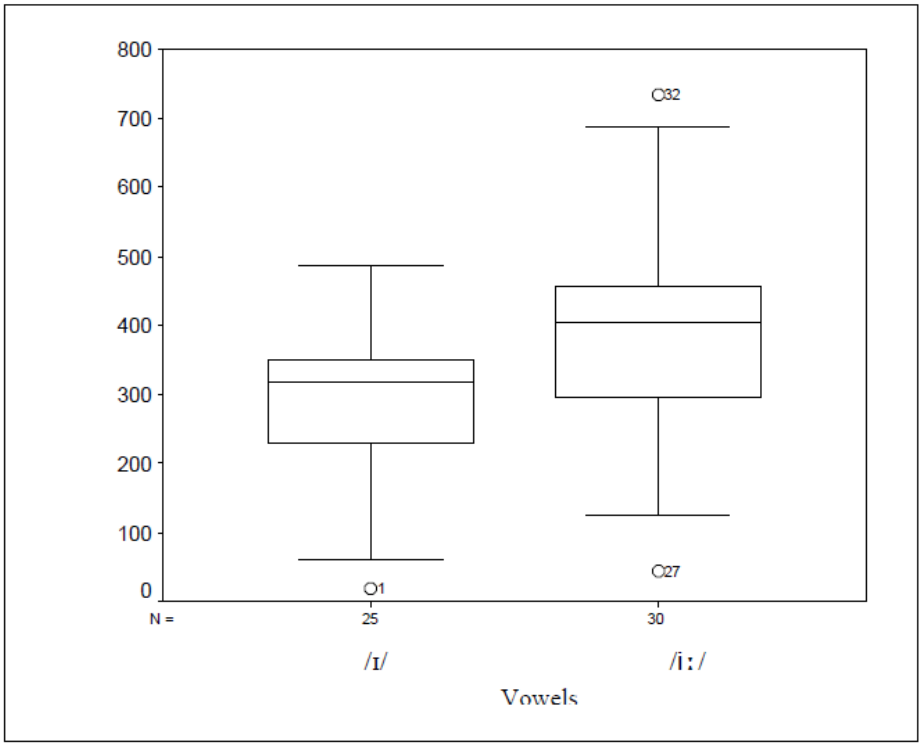


Figure 2.5: Vowel Chart for Malaysian English Vowels

(Sauce: Pillai, Zuraidah, Knowles & Tang, 2010)

The findings from this study indicate that "the contrast between vowels in Malaysian English does not match the classical notion of phonemic contrast". This lack of contrast in vowel quality was also found by Ahmad (2005), who examined the sets of vowels /ɪ/, /iː/ and /ʊ/, /uː/, and reported a similar length of these two vowel pairs. Figure 2.6 shows the summary of the vowel durations for these two pairs of vowels.



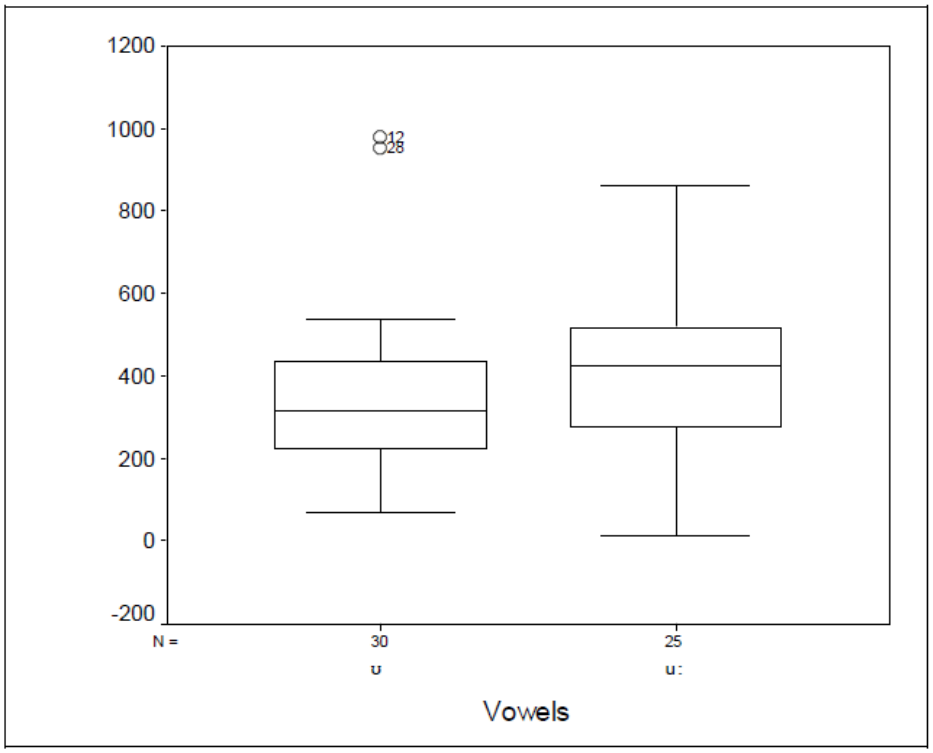


Figure 2.6: Box Plot Comparing Duration of /ɪ/, /iː/ and /ʊ/, /uː/

(Sauce: Wan Ahmad, 2005, p. 14)

As can be seen in the diagram, all the vowel pairs display overlapping distributions although the long vowel has longer range. This indicates that Malay speakers of English did not differentiate between long and short vowels. Zuraidah (1997) examined the Malay speakers production of ME vowels and discovered similar results that these speakers tended to produce /i/, / iː/ and /ʊ/, /uː/ with same quality and duration. For the /i/, / iː/ pair specifically, she reported that the subjects "seem to produce another realization whose quality correspond roughly to RP / iː/ but with less marked lengthening" (Zuraidah, 1997, p. 38). Thus, it can be assumed that there is a tendency that Malay speakers of English perceive no length contrasts in other's English production of vowel pairs.

2.4 Related Studies of Speech Perception

The scientific study of the perception of the spoken language has been a prolific and productive area in the field of phonetics and phonology. A myriad of research has been carried out assessing and exploring issues in language perception from various perspectives of language learning beginning in the 1950s (Jusczyk & Luce, 2002). Some draw the focus on ability of monolingual speakers with normal or impaired hearing identifying and distinguishing between the individual sounds of the native language as applied to the study of word recognition (Fowler, 1986; Fowler & Rosenblum, 1991); some concentrate exclusively on the ability of monolingual native speakers in the production of L2 speakers (e.g., Rogers, 1997); others highlight those aspects in perceiving and understanding a second or foreign language (Wang, 1997), whilst the rest deal with comparisons between both L1 and L2 (e.g., Polka & Bohn,

1996). However, these subsectors always converge on the assumption that production is contingent upon the perception that fluent and grammar-normative production requires a good perceptive basis (e.g., Best & Strange, 1992; Ioup, 1995). Flege (1995, p. 239) also stated in his work that “sounds in L1 and L2 are related perceptually to one another at a position-sensitive allophonic level rather than at a more abstract phonemic level.”

The beginning of the process of learning a second language is somewhat similar to that of the first language acquisition, where young children have to go through the stage of perceiving the sounds of L1 before proceeding to the next stage of production of them (Wang, 1997). However, concerning the ability of infants and adults of discriminating or perceiving the difference between non-native sounds, Werker and Polka (1993) state that infants have perceptual abilities to discriminate phonetic contrasts on a broad, language-universal basis, whilst, adults generally show a decreased ability in the discrimination of non-native phonetic contrasts in order to facilitate or enhance the perception of native phonetic contrasts. There are constant indications and proofs from research examining speech perception that listeners always have problems in receiving and understanding the production of second language speakers or learners. As suggested by Guion and Flege (1987) “adults who learn a second language (L2) perceive L2 sounds differently than monolingual native speakers of the target L2 do”. In this case, many theories have been proposed to explain why L2 learners always experience difficulty in perceiving phonemic contrasts in the target language (Lado, 1957; Flege, 1988; Best, 1994; Bohn and Flege, 1990).

Parallel researches have also been conducted comparing the perception of bilingual and monolingual speakers. Findings have shown that the vowel perception of late bilinguals often differs from that of monolingual natives (Flege, MacKay & Meador, 1999). Parallel findings were reported by Mack (1989) and Pallier *et al.* (1997) that early bilinguals may perceive L2 vowels differently than monolingual native speakers. However, the magnitude of native versus non-native differences seems to depend on the degree of perceived similarity of the L2 vowels to the closest L1 category (Polka, 1995; Best, Faber & Levitt, 1996). In another study, Flege (1991) examined the perception of monolingual Spanish speakers and Spanish speakers of English as their second language. Both groups were presented with English vowels /ɪ/, /i/, /æ/, /ɛ/ in the target words *bit, beat, bat, bet* as produced by native English speakers. They were asked to label the vowels heard in these words by circling one of the Spanish vowels /i/, /a/, /e/, /o/, /u/, or "none". It turned out that the second group, the bilingual speakers, was better in identifying vowels using "none" correctly more frequently than the monolingual speakers in a range of 42% to 18%. Mack (1989) also stated in his research that the perception of English vowels of monolinguals differed from that of early bilinguals; they approximated but were not identical. In 1994, Flege, Munro and Fox conducted an investigation of English monolinguals and native Spanish speakers of English, asking them to "rate the dissimilarity of tokens of two Spanish vowel categories, two English vowel categories, or one Spanish and one English vowel category". The performance of dissimilarity ratings increased with an increase in an F1-F2 acoustic space for both the native English and Spanish subjects who had mastered English as an L2; and in the

oddity discrimination task the result of monolingual native English subjects was better than the other group of subjects for adjacent triads.

Regarding this phenomenon in the learning of a second/foreign language, the importance of correct-perception towards the L2 contrast is a critical issue. There is a broad consensus that whether learners of a second/foreign language can achieve a native-like way of speaking is closely related to how well they perceive the sounds of the L2 that they are able to distinguish in the differences between the L1 and L2 inventories. This has been certified by the study of Yamada and Tohkura (1991), non-native speakers' rate of learning and their ultimate degree of success may be influenced by the perceived similarity of English consonants to consonants in the L1. Lado (1957) also found that if a segment only exists in the L2 but not in the L1, the learner will have trouble perceiving and producing the new segment. For example, several researches have shown that Japanese speakers generally have difficulty differentiating the English consonants /r/ and /l/, where there is a lack of these two sounds in their native language, even after years of experience with English (Logan, Lively & Pisoni, 1990). One explanation of this is that L2 learners or bilingual speakers tend to perceive the sounds of the L2 in terms of L1 phonology (Best, 1994; Flege, 1995). Both the phonological structure and phonetic characteristics of a speaker's native language will influence his perception of sounds in a learned foreign language. Whereas, it is not necessarily true that perceptual mastery of an L2 contrast always precedes the learners' ability to produce contrasting phonemes (Blankenship, 1991). She examined the English perception and production of Spanish bilinguals and discovered that even

though they could not perceive all the differences in the L2 vowels, they were still able to produce all the necessary L2 vowel contrasts in both the reading and interview tasks.

There are several factors that may affect perception of sounds of an L2; one of the mostly important is the age of acquisition of the L2. It is believed among many researchers that exposure to an L2 provides an advantage to L2 learners (Singleton, 1995). According to the critical period hypothesis (CPH) "foreign accent cannot be overcome easily after puberty" (Lenneberg, 1967: 176). It predicts that once a neurologically based critical period has been passed, the sounds of an L2 cannot be learnt perfectly (Patkowski, 1990; Long, 1990). This may be due to a decrease of the overall ability of perceiving the L2 sounds and being able to discriminate them from the L1 category. In Long's (1990) studies, he addressed the effect of age on second language phonological learning, concluding that a native-like pronunciation is impossible for many individuals if their first exposure is not before the age of six and for the remainder by about the age of 12. However, other researchers have found that the ability to discriminate between non-native sound contrasts declines by 4 years of age (Werker & Tees, 1983), whereas others have found that this ability may be lost at the age of 6 to 7 years (DeKeyser, 2000).

The second factor addresses the importance of experience in the learning of a second language. According to Perceptual Assimilation Model (Best and Strange, 1992), it is anticipated that L2 vowel discriminability may improve as a function of L2 experience. Hillenbrand and Gayvert (1993) investigated the ability of native English speakers in

identifying some English vowels, namely /ɪ/, /i/, /æ/, /ɛ/, /u/, /ʊ/, /ʌ/, /ɔ/, /ɑ/. The findings showed that the participants who were phonetically trained prior to the experiment bore quite encouraging performance with the greatest accuracy in the identification of the sound /i/ at a percentage of 95%, followed by the vowel /ʊ/ (85%), /ʌ/ (79%), /ɪ/ (77%), /u/ (73%), /ɔ/ (72%), /æ/ (64%), /ɛ/ (61%), and /ɑ/ (51%), though certain vowels were frequently confused with one another in a steady-state synthesis. It is also reported that the distance between F1 and F2 played a major role in the participants' performance. The greater the distance, the more accurate their identification. For example, they tended to replace /ɔ/ for the target /ɑ/; the vowel /ɪ/was easily mistaken as /i/ and /ɛ/ for the target /æ/; whilst all participants performed well on the discrimination of vowels /i/ and /u/. Similarly, Wells and Colson (1997: 97) found that phonetically naive people tend to hear all vocoids as related to the vowel sounds phonemic in their language. However, other researchers have been less quick to certify the status of experience in sound perception. Flege, Munro and Fox (1994) declared, "the dissimilarity ratings of experienced and inexperienced Spanish subjects did not differ significantly".

The effect of native language experience on the perception of non-native language phones has been treated in terms of developmental changes. For adult L2 learners, language specific experience has proven to have an effect on the perception of non-native speech by the end of the first year in life (Werker and Polka, 1993). However, the effect of it is "neither absolute nor permanent" (Best, 1994). It is worth

noting that studies have also reported that adults can discriminate some non-native phonetic distinctions without explicit training (cf. Flege and Port, 1981).

2.5 Discrimination and Determination of Vowel Sounds

Since both native and nonnative speakers of English face difficulties in discriminating among certain English vowels (Peterson & Barney, 1952; Flege, 1991; Wheeler, 1995), discrimination and identification are better for vowels that are more dissimilar from one another (Flege, Munro & Fox, 1994). According to Borden, Harris & Raphael (1994), the primary acoustic cues for discriminating and identifying vowels consists in their respective formant frequencies. Acoustically, vowels are characterized by the formants, especially the first two (F1 and F2). Mack (1989) examined the identification and discrimination of the members of a synthetic English /i/-/ɪ/ continuum by English monolinguals and early bilinguals manipulating both English and French. He found that there was no noticeable gap in their abilities of discriminating vowels, though they did differ in the location of their phoneme boundaries. The difficulty in discriminating certain vowels derives from their closer proximity to one another within this F1-F2 acoustic space (Bond & Robey, 1983; Hillenbrand & Guyvert, 1993; Peterson & Barney, 1952). Results of empirical studies based on the adoption of these methods could be used to determine the phonemic contrasts between the L1 and L2 that required special instruction and training depending on the relative level of difficulty of perceived difference.