The Perception and Production of SSBE vowels by Syrian Arabic learners: The Foreign Language Model

Rana Alhussein Almbark

This thesis is submitted for the degree of

Doctor of Philosophy

Department of Language and Linguistic Science

September 2012

Abstract

This thesis presents an examination of the perception and production of Standard Southern British English (SSBE) vowels by Syrian Arabic (SA) Foreign language (FL) learners. The focus of this thesis is the FL learners who learned their English in their country and mostly by non-native teachers. Thus, by definition, the FL learners do not have native English input on a daily basis. This thesis reports on an empirical investigation of the Second Language (L2) perceptual and production patterns of a group of FL learners, which has received little interest in the literature, combining insights from current cross-language speech perception (Perceptual Assimilation Model) (Best 1994, 1995, 1999) and L2 learning models (Speech Learning Model) (Flege 1995). These models were mainly developed to account for early and advanced L2 learners, respectively. Thus, this study aims to develop an account for the perception and production of FL learners based on current L2 models.

Results indicate that the specific learning context of FL learners is reflected in their perception and production patterns. For example, these learners live in a predominantly L1 environment, and their L2 input is mainly taken in a classroom and mostly by local teachers. However, this study argues that though FL learners lack native L2 input, they do have access to the phonology, syntax, and structures of the L2 via direct teaching. It is also shown that the perceptual patterns of the learners succeeded in predicting their production patterns, which has implications on the perception-production link for L2 learners, in general, and for FL learners in particular.

The main outcome of the present thesis is that it develops an account of the perception and production of FL learners. It outlines the main principles for a proposed Foreign Language Model, in which the peculiarities of FL learners are taken into consideration compared to other groups of learners.

Content

List of Tables

List of Figures

Abbreviations

Abbreviation	Full text
AE	American English
AOA	Age of Arrival
AOL	Age of Learning
CG	Category Goodness
DP	Distinct acoustic Parameters
DRT	Direct Realist Theory
FL	Foreign Language
L1	First Language (Syrian Arabic)
L2	Second Language (Syrian Arabic learners of English)
LOR	Length Of Residence
MT	Motor Theory
NE	Native English (SSBE participants)
NLM	Native Language Model
OP	Overlapping acoustic Parameters
PAM	Perceptual Assimilation Model
PAT	Perceptual Assimilation Task
SA	Syrian Arabic
SC	Single Category
SSBE	Standard Southern British English
TC	Two Categories
UC	Uncategorised- Categorised
VOT	Voice Onset Time

Acknowledgments

Looking back, I realise how thankful and grateful I am to all those people around me, who helped me and supported me throughout the ups and downs of writing my PhD. First, I would like to thank my supervisor Sam Hellmuth for her outstanding guidance and academic support. I would like to thank her for the countless hours that she spent with me and for me.

I also would like to thank the members of my Thesis Advisory Panel Paul Foulkes and Danijela Trenkic for their valuable suggestions and comments which I got throughout my PhD. I also thank Dom Watt for his valuable feedback and suggestions on vowel analysis and normalisation. I am very grateful to Tamar Keren-Portnoy for coming to our research groups through which I continued to receive her informative comments and suggestions. I would like to express my sincere gratitude to Ghada Khattab and Jalal Al-tamimi, who provided me with the advice, resources and references. I am very grateful to the statistician Arthur Ricky, who helped me run and understand the statistical analyses I needed for my thesis. I thank Huw Llewelyn-Jones who provided me with all the technical support and advice during data collection.

I am mostly grateful to Asia institute for languages in Damascus, which made this thesis possible, by giving me permission to record some of their students and by allowing me to use their class rooms to do my recordings. I would like to thank all of my Syrian and British participants who were very kind to participate in my study.

I would like to thank all of my friends and colleagues who supported me and helped me throughout the hard and happy times. I would like to thank Becky Taylor whose suggestions and comments were always helpful. I am very grateful to her for giving her time and effort to proof-read my thesis. Thanks to Ghazi Algethami for his endless support and informative comments.

Author’s Declaration

This thesis has not previously been accepted for any degree and is not being concurrently submitted in candidature for any degree other than Doctor of Philosophy of the University of York. This thesis is the result of my own investigations, except where otherwise stated. Explicit references acknowledge any other sources used in this thesis. I declare that part of the results of my thesis was presented in these conferences:

BAAP 2012

Postgraduate conference on Arabic Phonetics and Phonology PCAPP 2011

Accents 2010

CamLing 2010

NewSounds 2010

My paper in NewSounds 2010 has been accepted and published in a peer-reviewed proceeding: Almbark, R. A. 2011. Production and Perception of SSBE Vowels by Syrian Arabic Speakers. In *Achievements and in Perspectives SLA of Speech: New Sounds 2010*, ed. Magdalena Wrembel, Małgorzata Kul and K. Dziubalska-Kołaczyk, 15-26. Peter Lang: International Verlag der Wissenschaften.

I hereby give consent for my thesis, if accepted, to be made available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed . . . . (candidate)

Date . . . .

1 Introduction

The speech of Second Language (L2) learners has been the target of a great body of research during the last few decades. The main interest of L2 research has been L2 learners who acquired the L2 in natural settings over a considerable time, i.e. immigrants (Flege 1995, Flege and Mackay 2010). Another group of L2 learners, which has been extensively researched, is foreign language (FL) learners who have little or no L2 experience (Best 1994, 1995, 1999). A third group of learners is FL learners who were taught their L2 in their home country, mostly by local teachers. The latter group has been extensively researched within an educational context and from a student-teacher perspective. However, the speech of the third group has received little attention in L2 perception and production research despite the greater number of its members compared to the previous two groups. For example, English is the most common target language for teaching all over the world. So, how do we account for the perception and production patterns of the speech of millions of FL learners of English? How can the available L2 body of research account for the speech of those learners?

The present thesis advocates a theoretical approach to investigate the perception and production of FL learners, an approach which is nonetheless firmly grounded in a combination of cross-language and L2 theories. The present thesis provides empirical evidence that examination of the speech of this group of learners requires a horizontal view of the current L2 theories as none of these theories account for the speech of FL learners. Additionally, none of these theories explicitly account for the perception-production link of L2 learners' speech.

Therefore, the present thesis employs some insights of the current L2 theories: Native Language Magnet Theory (NLM), Perceptual Assimilation Model (PAM), and Speech Learning Model (SLM), to investigate the production and perception of the Standard Southern British English (SSBE) vowels by Syrian Arabic (SA) FL learners. Exploiting these theories will enable us to decide which one explains the results of the present thesis best, and to decide which aspects of these theories can be modified to be extended to our target learners.

This chapter is organised as follows: in this section, an overview of the general topic of the present thesis is presented. The motivations and research questions behind this research are presented in section 1.1. Then, the overall structure of the rest of the thesis is presented in section 1.2.

1.1 Motivations and Research Questions

This thesis is an attempt to examine the perception and production patterns of SA learners of English. This research aims to fill a gap in the literature in two ways. First, Standard Arabic is well known as a 3-quality vowel system (Mitchell 1993, Newman and Verhoeven 2002) with short/long distinctions. However, little research has been conducted to examine Arabic vowels experimentally, in particular dialectal Arabic vowel systems. The previous research on SA vocalic system is incomplete, as some researchers used only one speaker or only auditory analysis or only analysis of the basic contrasts (Abou Haidar 1994, Allatif 2008, Allatif and Abry 2004, Cowell 1964). Thus, the present study contributes to the literature on Arabic by providing an acoustic description of the full SA vowel system, which has not been done in any previous research. Providing this description is also vital to the present thesis since SA is the First Language (L1) of the target learners of this study.

Second, English is a universal language and learning it has become compulsory in most (if not all) of the Education systems of the Arab countries. SA is one of those systems in which teaching English starts from the first grade (Scobbie 2007). Therefore, the present thesis contributes to the L2 research by examining both the perceptual and production patterns of English vowels by SA FL learners compared to Native English (NE) and L1 patterns, for the first time.

An advantage of this thesis is the combination of L2 speech perception and production. Examining both aspects of L2 speech helps in relating the participants' output to their own perceptual patterns. For example, the vowel systems of SA and British English are different from each other qualitatively and quantitatively, and examining how SA learners, with their limited vowels, manage to perceive and produce a rich vowel system of a language such as English will add more to our understanding of the production-perception link. This study will add to our understanding of how/whether L2 speech perception affects L2 speech production and vice versa.

In sum, the present thesis aimed to answer the following questions in the analytic chapters of the thesis:

- To what extent do current L2 models (NLM, PAM, and SLM) account for the perceptual and production patterns of SA FL learners?
- Which of these models can be extended to account for the target learners of the present thesis? And what are the required modifications, if any, to be extended?
- What does the present thesis tell us about the perception-production link in L2 speech compared to that in L1 speech? In light of the results of the present thesis, what can we infer about the perception-production link for the FL learners compared to advanced L2 learners and the L1?

1.2 Structure of the Thesis

The present thesis is organised as follows. Chapter 2 provides a background review of the mechanisms and theories of speech perception and speech production, with the focus on vocalic sounds. It also provides a background review of the theories of second language learning including the Native Language Magnet Theory (NLM), Perceptual Assimilation Model (PAM), and Speech Learning Model (SLM). This chapter also presents a detailed review of the literature on the vowel systems of the First and Second Languages under investigation (L1: Arabic with a focus on SA, L2: English with a focus on SSBE). The background chapter concludes with a review of some of the L2 studies involving the two languages; and in particular the studies which employ the predictions of the L2 models under investigation.

Chapter 3 presents the design of perceptual and production experiments which were conducted for a pilot study to examine the SA learners' perception and production of all English vowels. Two perceptual experiments: a Perceptual Assimilation Task (PAT) and an Identification task, and a production task were run in the pilot study. The results and findings of the pilot study were used to direct the focus of the main study in the following chapters (Chapters 4-7) to particular English vowels with specific predictions and research questions, as well as to identify areas of difficulties and flaws in terms of the research design.

The main study begins with Chapter 4 which presents a detailed acoustic analysis of the participants’ L1/SA vowel system including its identified phonological and allophonic categories. This chapter had a great effect on the rest of the thesis, as it was possible to make stronger claims and predictions about L1 effect on the participants’ L2 perceptual and production patterns in the following chapters.

Chapter 5 presents the first perceptual experiment in the main study of this thesis, i.e. the Perceptual Assimilation Task (PAT). The PAM L2 model was mainly used to predict the perceptual assimilation patterns of the SA participants, which were also based on the findings of the pilot study in Chapter 3 and the SA vowel system in Chapter 4. Chapter 5 includes a presentation of the methods, results, and discussion of the main findings of this task. Based on the main findings of Chapter 3 and 4, and the PAT in Chapter 5, specific hypotheses and predictions were examined in the second perceptual experiment, the identification task, in Chapter 6. Similarly to the previous chapter, Chapter 6 included the methods, results and discussion of the main findings. After that, Chapter 7 presents the production task which had specific hypotheses and predictions based on the perceptual experiments in Chapters 5 and 6 and the SA vowel system in Chapter 4. Similarly, Chapter 7 spells out the methods, results, and discussion of the main findings.

Finally, Chapter 8 presents a general discussion of the main findings of the present thesis. It also discusses the main research questions in section 1.1 above based on the results and the findings of the present thesis. This chapter concludes with a summary of the main conclusions and some suggestions for future work.

2 Background

This chapter provides a theoretical and practical background overview of the previous literature which helped in developing the research questions of this study. The structure of the chapter is as follows: Sections 2.1 presents a review of some theories of speech perception in first language (L1) acquisition, with a special focus on the Motor Theory (MT) and the Direct Realist theory (DRT) of speech perception. An overview of the nature of speech perception and production processes and their developments in first language acquisition will set the stage for the subsequent discussions of second/foreign language perception and production. Section 2.6 presents a review of the only and widely used speech production theory, i.e. the source-filter theory, with a special focus on the acoustic analysis of vowels, which is the main interest of the present thesis.

Section 2.2 presents an overview of second language (L2) theories. The perception and production of L2 sounds have been discussed within the framework of the Native Language Magnet theory (NLM), Perceptual Assimilation Model (PAM), and Speech Learning Model (SLM). Then, the claims and predictions of these models are compared in Section 2.2.4.

Section 2.3 presents a review of the vowel system of Arabic, i.e. the L1 in this study, including vowel quality, vowel quantity, and diphthongs. An overview of the vowel system of Dialectal Syrian Arabic, which is the main focus of this thesis, is presented in Section 2.3.4. (Similar to Arabic vowels) Section 2.4 provides a review of the vowel system of Standard Southern British English (SSBE), which is the L2, including vowel quality, vowel quantity, and diphthongs. Section 2.5 presents a review of the L2 studies which have been conducted on Arabic learners of English, in particular. Finally, a summary of this chapter is given in Section 2.7.

2.1 Speech Perception

One of the main interests of the present thesis is to examine L2 speech perception. In this section an overview of speech perception theories in first language L1 is presented, which will set out the discussion of second/foreign language speech perception. The term *perception* is associated with many events including speech. These events can be direct sensory actions (e.g. vision and listening) or indirect mental processes (e.g. thinking and comprehension) (Diehl et al. 2004, Fowler and Dekle 1991). The Merriam-Webster dictionary[1] defines the word *perception* as ‘a result of perceiving: observation’, ‘a mental image: concept’, ‘awareness of the elements of environment through physical sensation: colour perception’, ‘physical sensation interpreted in the light of experience’, ‘quick, acute, and intuitive cognition’, or ‘a capacity for comprehension’ (James M Scobbie et al. 2008). These definitions associate *perception* with direct sensory actions such as environment and physical sensation. Perception is also associated with mental processes such as concept, awareness, experience, and cognition. Thus, any definition or account of speech perception should have one or both associations: sensory actions and indirect mental processes.

Research in speech perception aims to answer the questions of how listeners recognise, interpret and understand speech. When language is produced, it is assumed that the acoustic signal of the sound being processed is mapped onto a more abstract form, whether this form refers to phonemes, syllables, or features (Casserly and Pisoni 2010, Fowler and Dekle 1991, Liberman et al. 1957). This mapping process has been the subject of research at the Haskins Laboratories (Liberman, et al. 1957). It has been referred to as categorical perception, which was an important finding at the. It has been argued that categorical perception can be used to explain, to some extent, the mismatch between the physical world (represented by the acoustic signal) and the cognitive world of speech (represented by the abstract phoneme categories), which has been one of the fundamental issues in speech perception research (Casserly and Pisoni 2010).

The perception of the acoustic signal and mapping it to a particular phoneme is not a straightforward process. The acoustic signal varies for a single phoneme due to several

[1] The definitions were extracted from Merriam-Webster online dictionary (http://www.merriam-webster.com/dictionary/perception) on 26/02/2011.

factors such as neighbouring contexts and speech rate. For example, Liberman and his colleagues (1957) used a series of synthetic CV syllables which differed in the extent and direction of second formant (F2) transitions. Varying F2 transitions were used as an acoustic cue to various initial consonants /b d g/. Listeners were asked to identify the initial consonants of the /bV/-/dV/-/gV/ syllables, and they were also asked to discriminate closely located pairs of stimuli over the synthetic continuum. The results showed that there were discrete boundaries in the identification between the three phoneme categories, and that discrimination across these boundaries was significantly better than discrimination within a phoneme category.

An implication of categorical perception is that speakers of different languages have different phonemic boundaries for similar categories. Thus, it is argued that listeners of the same language should not have a problem with variation in the acoustic signal, and that they should be able to divide the speech signal into similar categories, regardless of variation and overlap in the acoustic cues of speech sounds. For example, languages like English and Spanish were found to have different phonetic Voice Onset Time VOT values to represent +/- voiced stops. English +voiced stops are realised with short positive VOT and –voiced stops are realised with long positive VOT. On the other hand, Spanish +voiced stops are realised with long negative VOT and –voiced stops are realised with short positive VOT. Using synthesised stimuli, Lisker and Abramson (1970) asked the participants from both languages to identify stops in the stimuli using orthography of their languages. The results showed that timing of the VOT had a great effect on boundaries locations identified by the participants from both languages. Therefore, one can argue that the categorical perception of VOT depends on language background, where listeners become more sensitive to particular phonetic differences rather than others, and these differences appear to play a distinctive role in their language (Diehl, et al. 2004).

There is a great amount of literature on speech perception, which explores the perception of speech from different angles: articulatory, acoustic/auditory, or cognitive. Thus, a number of theories have been suggested to explore how speech is perceived. The main focus of these theories was to determine the nature of the object of speech perception input: acoustic or gestural, and whether speech perception requires a specialised mechanism to decode the speech signal or not. Among the earliest theories of speech perception are Motor Theory (MT) (Liberman et al. 1967, Liberman, et al. 1957, Liberman and Mattingly 1985) and Direct Realist Theory (DRT) (Fowler 1980, 1981, 1986, 1996, Fowler and Dekle 1991), which can be grouped together as gestural theories since both assume that the objects of speech perception are articulatory gestures. The main L2 model used in this thesis is built on a gestural based theory; therefore, the principles of these two theories will be compared and discussed in detail in the following sections 2.1.1 and 2.1.2.

On the other hand, the acoustic landmarks and distinctive features model which was proposed by Stevens (Stevens 1972, Stevens 2002, Stevens and Keyser 2010), is considered an auditory model. According to this model, speech perception begins by identifying specific acoustic landmarks (such as higher frequencies of vowels or discontinuity of consonants) in the acoustic signal. Then, these acoustic landmarks are used to uncover the features intended by the speaker. The distinctive features referred to by Stevens represent features which can be used contrastively in a language, for example [+/- rounded], [+/- back], [+/- high] for vowels; [+/- sonorant], [+/- voiced], or [+/- nasal] for consonants (Stevens 2002). It is assumed that each distinctive feature is defined by a set of articulatory actions, and these actions correspond to particular articulatory correlates (Stevens 2002). Finally, these features are mapped into the lexicon.

The acoustic landmarks and distinctive features model is based on the quantal theory of speech, which assumes a nonlinear relationship between articulation and corresponding acoustic signal. Consequently, small changes in the articulation may not change a category and therefore our auditory system might not respond to small changes within the articulatory dimension of a particular feature. Therefore, this model gives equal status to the articulatory and acoustic dimensions in speech perception, and unlike other models, it attempts to map the phonetic and the phonological levels of the language (Stevens and Keyser 2010).

A more cognitive model of speech perception is the Perceptual Magnet theory which is interested in the mental representation of speech categories. This theory was proposed by Kuhl (1994) and was based on the Native Language Model (NLM). The NLM describes the development of the phonetic categories of infants' native language as a result of experience with the ambient language. Perceptual Magnet theory assumes that during the early stages, infants are able to differentiate all possible human sounds using their ‘general auditory mechanisms’. As a result of experience with the native language, infants’ sensitivity to the phonetic representations of their ambient language increases; and these representations constitute ‘prototypes’ which function as perceptual magnets for the members of their categories. Thus, the perceptual magnet effect is supposed to structure infants’ phonological space (Kuhl et al. 2008).

In contrast, for adult second or foreign language learners the perceptual magnet effect reduces adults’ sensitivity to foreign categories which are phonetically close to their native prototypes (Iverson and Kuhl 1995). An implication of the perceptual magnet effect would be that it is difficult for adults to perceive phonetic contrasts that do not constitute contrastive phonological categories in their native language, which has been an important issue in second language research. The NLM effect on second/foreign speech perception is discussed in more detail in section 2.2.1 below.

The perceptual results of the present thesis are analysed and discussed within the framework of the Perceptual Assimilation Model (PAM), which is based on a gestural view of speech perception. Thus, the gestural theories mentioned above, i.e. Motor Theory (MT) and Direct Realist Theory (DRT), will be discussed and compared in detail in the following sections.

2.1.1 Motor Theory (MT)

The Motor Theory (MT) of speech perception was developed by Liberman and his colleagues (Liberman, et al. 1967, Liberman, et al. 1957, Liberman and Mattingly 1985). Diehl et al. (2004) summarised the basic assumptions of MT as follows:

- Articulatory events constitute the objects of speech rather than acoustic or auditory events.
- Human listeners recover the articulatory events by neuromotor commands of the intended gestures rather than the actual articulatory movements or gestures.
- Speech is perceived by an innately specialised mechanism or decoder which is unique to speech and to humans.

- The mismatch between the phonemes (invariants) and acoustic signals (variants) is claimed to result from the coarticulation of consonants and vowels which results in a complex mapping between phonemes and vocal tract shapes.

The MT of speech perception is a gestural theory, as opposed to acoustic or auditory theories, since it assumes that the articulatory gestures are the object of speech perception. The MT postulates that speech is perceived by reference to how it is produced; that is, listeners use their own knowledge of how phonemes are articulated when perceiving speech. Listeners are believed to retrieve the phonetic information from the articulatory gestures such as lip rounding or tongue backing using a specialised mechanism. Thus, according to this theory, what is perceived is the intended rather than the actual articulatory gestures.

In a revised version of the MT, Liberman et al. (1985) claimed that speech perception and speech production are intimately linked since they share the same set of invariants (i.e. the articulatory gestures used in production and the intended gestures predicted in perception). They believed that this link is innately specified and not learned, which suggests that humans should have inherent knowledge of the speech production mechanism. However, one can argue that infants are not likely to have such inherent knowledge of speech production mechanism because they spend their early months or even years learning to control and use their vocal tract to become more like adults. In favour of this argument, Kuhl et al. (1996) suggest that two main factors contribute to changes in child vocalisations which are ‘anatomical restructuring’ and ‘vocal learning’. Anatomically, the child’s vocal tract changes during the first year of life to be like that of adults. Accordingly, the child’s ability to control his/her articulators increases. Also, infants attempt to acquire and learn their language specific properties via imitation, which have been suggested to take place between 10 and 12 months of age (Kuhl and Meltzoff 1996).

The finding of categorical perception, i.e. showing discrete boundaries between phoneme categories, (Liberman, et al. 1957) is considered as an important characteristic supporting MT, which predicts that due to categorical perception listeners can discriminate two stimuli across phoneme boundaries, but they are unlikely to discriminate stimuli within the same phoneme category. Following this prediction, it is argued that MT can account for the variation in the acoustic signal. That is, the ways that phonemes are produced and perceived, using articulatory gestures, are claimed to show more consistency than the ways they are acoustically represented and perceived.

Nevertheless, one may claim that variation in the acoustic signal is inherently present and inevitable. The presence of this variation can be due to several factors linguistic (e.g. surrounding context and speech rate) and/or extra-linguistic (e.g. vocal tract anatomical differences, dialect or language). Thus, any attempt to explain speech perception should take into account this variation and not try to ignore or eliminate any variation. In sociolinguistic research, listeners were found to use such variation in their communication. For example, Niedzielski (1999) found that providing social information such as 'speakers' nationality' to half of her participants caused variation in the responses between the two groups of participants. Even though the articulatory gestures were the same, an extra-linguistic cue varied the results.

Later research in speech perception also showed that categorical perception is not restricted to phoneme perception; it has been found in babies who cannot talk. It has been found that even four month old infants display categorical perception (Eimas et al. 1987). It is not even restricted to humans. Kuhl et al. (1975) found that trained chinchillas demonstrated categorical perception of /t/ and /d/. The results of an identification task of synthetic stimuli varying in VOT showed that the phonetic boundaries for chinchillas and native English speakers were the same. These findings suggest that a more general auditory mechanism plays a role in categorical perception rather than a specialized mechanism (Diehl, et al. 2004). Thus, it is not always the case that we use the same neuromotor commands, of the articulatory gestures that we use when we speak, to interpret what we hear.

Generally then, MT has not been found sufficient to explain speech perception. Research in speech perception has provided evidence that speech perception does not need a mediate or a special mechanism. Instead, humans use a more general perceptual mechanism similar to the one used by non humans. Additionally, coarticulation is argued to be inherent in the speech signal and is not a case of mismatch between the phonemes and vocal tract shapes. It has been also suggested that the actual rather than the intended articulatory gestures are the objects of speech perception. Some of these arguments and suggestions have been proposed by the Direct Realist Theory (DRT), which is another gestural theory of speech perception. In the following section the principles of DRT are discussed and compared to those of the MT.

2.1.2 Direct Realist Theory (DRT)

In the 1980s, another gestural theory of speech perception was developed by Fowler (Fowler 1980, 1981, 1986, 1996, Fowler and Dekle 1991), which is known as the Direct Realist Theory (DRT). Similar to the MT, the DRT assumes that the objects of speech perception are the articulatory gestures. However, DRT suggests that the actual vocal tract movements are the objects of speech perception rather than the intended gestures as in the MT.

The DRT also contrasts with the MT regarding the presence of an innate specialised mechanism for speech perception, which is speech-and human-specific. Fowler argues that speech perception is, generally, similar to other physiological perceptual processes such as visual perception (Fowler and Dekle 1991). Fowler (2008) claims that perceiving the motor articulatory actions of speech is direct and analogous to perceiving a person walking when you see a person walking. However, one can argue that not all articulatory motor actions are visible to listeners. Moreover, coarticulated speech entails that some motor actions occur simultaneously which makes it even more demanding for listeners to perceive them. Nevertheless, listeners perceive coarticulated speech effortlessly. Furthermore, the DRT views coarticulation in the speech signal as a result of gestural coproduction of vowels and consonants rather than temporal overlap as suggested in the MT. For example, Fowler argues that vowels are produced continuously even during the production of a following or a preceding consonant (Fowler 1980, 1981, 1986).

In the formation of her theory, Fowler followed the principle of ‘direct perception or direct realism’ of the General Theory of Direct Perception by (Gibson 1966, 1979). The basic assumption of direct realism is that speech perception is direct and that listeners can use the acoustic signal to infer the actual articulatory gestures. Although the acoustic signal is produced as a result of combination of articulatory gestures, a given speech signal can be produced using several vocal tract configurations (Diehl, et al. 2004, Fowler 1981, 1996). Also, two phonological segments with different acoustic signals can be perceived as similar. Fowler (1996: 1731) claims that hearing the two phonological segments with different acoustic signals as similar can be seen as a consequence of producing them with similar invariant phonological gestures, which she defines as ‘the public actions of the vocal tract that cause structure in acoustic speech signals’.

According to the DRT, perceiving two categories as phonologically distinct denotes that the two categories are produced with different constellations of phonological gestures and thus with acoustically distinct parameters, i.e. such as duration and F1-F2 formants for vowels. However, perceiving two categories as phonologically similar does not necessarily denote that they are produced with overlapping acoustic parameters. The DRT suggests that phonologically similar categories are produced with similar set of phonological gestures which might result in distinct or overlapping parameters in the acoustic signal.

The gestural theories of speech perception, particularly DRT, had an influence on studies of the perception of second/foreign language speech. Fowler (1986) defines the phonetic segment as ‘coordinated gestures of vocal-tract structures’. Thus, arguably, the different gestural combinations used by speakers of different languages form the phonological space for these languages. Based on the DRT, Best argues that languages differ in ‘which gestures they combine’ and ‘ the particular phasings they set between the gestures’ (Best 1995: 193). One can claim that any difficulty in producing or perceiving second/foreign segments might stem from differences or similarities in the articulatory gestures between the native and the second language. This claim has been introduced and examined within the framework of second language theories, which will be discussed in Section 2.2.

Both MT and DRT agree on a relationship between speech production and speech perception. According to these gestural theories, speakers produce the articulatory gestures and listeners perceive the intended gestures (MT) or the actual gestures (DRT). Assuming this close relationship has ‘strong implications for understanding both the phonetic details and the phonological organisation of a language (Best 1995: 172)’. Thus, the different combinations of the articulatory gestures constitute the phonological categories for a particular language, which coincides with the principles of articulatory phonology, in which gestures are analysed as speech input whereas the corresponding acoustic signal is analysed as the output (Browman and Goldstein 1992).

The previous sections reviewed the gestural theories of speech perception. In order to understand people as listeners, acoustic data is needed. On the other hand, to understand people as speakers, articulatory data is needed (Scobbie et al. 2008). So, if we want to understand the perception-production link in terms of gestures, we need to know how the articulatory gestures suggested by MT and DRT map into acoustic outputs.

Acoustic analysis is argued to enable examination of speakers' output, but not the articulatory strategies of the speakers' outputs. Scobbie, et al. (2008) found that anterior covert gestures, though strong in derhotic speech, have negligible effects on the acoustic output of rhoticity and thus on perception (if the acoustic signal what is perceived by listeners). This leads to question the role of gestures in speech perception. If what people use in speech perception and production is representations of articulatory gestures, then it would not be possible for L2 learners to perceive actual articulations which do not fit their own gestural representations.

A language phonological system does not only include specifications of its phonemes but also description of its phonetics, allophones and phonotactics. How do gestural theories account for the allophonic and phonotactic variations of the language? Gestural theories (Fowler 1996) argue that the articulatory gestures used for phonological segments cause variations in the acoustic output due to coarticulation in the vocal tract gestures. Fowler Fowler (2008) claims that variations due to coarticulation should be exploited to recover the phonological gestures and not treated as noise to be tolerated. Similarly, Liberman and Mattingly (1985) suggest that an efficient perceptual process uses the acoustic output to recover the phonological gestures. Thus, the MT and DRT are mainly interested in segments/phonemes and treat coarticulation as a source of information to recover the gestures used to produce a particular segment rather than a source which produces discrete constellation of gestures in particular contexts.

In order to resolve the debate of the nature of perceptual representations, parallel articulatory, auditory, and acoustic analyses are needed. The present study does not aim to resolve this debate and how these analyses would be translated in production in the

L1 and accordingly in the L2. Therefore, this thesis adopts the Perceptual Assimilation Model, but not its assumption that speech perception is based on articulatory gestures.The following section aims to review theories of L2 learning which aimed to account for the perception and production patterns of L2 learners.

2.2 Second Language Learning Models

This section presents some contextual background on the main models of second/foreign language speech perception and production against which the research reported in this thesis is analysed and discussed.

Learning a second language is a complex process that involves mastering many layers: phonetic, phonological, semantic, syntactic and pragmatic. Arguably, for a learner to become native-like, it is important to master the different aspects of the language in a native-like fashion. The speech of adult learners of a second language is usually associated with different degrees of foreign accent. Munro (1993: 40) defines foreign accented speech as “speech which differs audibly from native-speaker norms”. Among these layers of language, it can be argued that the phonetic and phonological aspects distinguish the speech of non-natives from natives more than any other aspect. Therefore, having a foreign accent which exhibit segmental and suprasegmental errors is arguably a clearer marker of a speaker as non-native than syntactic or morphological errors, in particular, for foreign language learners who do not have direct access to the native L2 input; but do have access to the syntactic and lexical knowledge in the classroom.

Researchers have examined other factors that may play a role in the perceived accent, such as the fluency and speaking rate of L2 speech. Usually, L2 speech includes pauses, hesitation and many lexical errors, which make non-native speech sound slower than the native’s, particularly for low-proficiency learners (Wood 2004). Although fluency has been found to play a role in perceived accentedness, current L2 research suggests that it is mainly based on linguistic factors such as the segments and the prosody rather than on the fluency of L2 speech (Best 1995, Derwing et al. 2004, Flege 1995). Overall, foreign accent can be attributed to all of the different linguistic and extra-linguistic factors examined in the literature. The findings of the studies examining foreign accented speech suggest that one of these factors may contribute to the perceived accentedness more than the others.

During the last few decades, several theories of second language (L2) acquisition have been suggested to give a plausible account for the perceived heavy accentedness in the speech of adult learners. The main goal of these theories is to explain pronunciation errors in the speech of second language learners. Among the earliest theories is the Contrastive Analysis Hypothesis (CAH), which was first proposed by Lado (1957), it assumes that native language interference causes speakers' deviation from the target language. Lado (1957: 2) claims that the main source of difficulty that adult learners encounter is the difference between the native language (L1) and the target language (L2): "Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult".

Lado's main argument, which is still assumed by most L2 theories, is that comparing the phonological elements of the learners' L1 and that of the L2 would be able to predict the errors encountered (Davidson 2011). Under the effect of L1 transfer the CAH assumes that the way L2 phones are perceived affect the way they are produced. Based on this argument, the errors in learners' L2 speech should have a plausible explanation related to L1 transfer. Research in L2 acquisition provides many examples which coincide with this assumption. For example, it is well documented that adult Japanese learners of English have difficulty perceiving and producing English /ɹ/ and /l/ due to the assimilation of these liquids into a single Japanese category /ɾ/, where English /ɹ/ was perceived and produced better than /l/ (Aoyama et al. 2004, Flegea et al. 1996).

Following L2 research has shown that L2 acquisition is not as simple as the CAH suggests (Davidson 2011). The relationship between production and perception is not an absolute one. For a particular language, some L2 structures might be easier to perceive than to produce, or vice versa, which suggests that perception and production are two different processes (Klein 1986: 26). For example, the English /b/-/p/ distinction is predicted to be hard to perceive by Arab learners of English due to the lack of the voiceless /p/ stop in Arabic. However, Flege et al. (1981) found that Saudi learners were able to detect the phonological distinction of English /b/-/p/, which is similar to that of /d/-/t/ and /g/-/k/. However, producing this distinction is not always easy for Arab learners. From an articulatory perspective, it is difficult for Arab speakers to stop voicing during the production of English /p/. This difficulty arises from L1 transfer but it does not affect perception as much as the articulatory settings during production. It might be easier for L2 speakers to produce /p/ as instances of their L1 /b/ than giving more effort to change their articulatory settings to produce /p/.

The analysis of the CAH was mostly descriptive and its main focus was L1 transfer, which has been argued to be insufficient to explain L2 perception and production patterns. Subsequent L2 theories have been formulated to account for the role of L1 transfer in L2 phones learning in terms of perception and/or production. In the following sections, three major models of second language acquisition, which contribute to the analysis and discussion of the results of this thesis, are presented in detail: Native Language Magnet theory (NLM), Perceptual Assimilation Model (PAM) and Speech Learning Model (SLM), followed by a comparison among these models.

2.2.1 Native Language Magnet Theory

As mentioned in section 2.1 above, the Native Language Magnet theory (NLM), which was proposed by Kuhl (1994), explores speech perception from a cognitive perspective. This theory assumes that, during the first year of life, infants build up phonetic representations for the categories of their native language, which are called ‘prototypes’ (Kuhl and Iverson 1995, Kuhl 1994, Kuhl, et al. 2008). Kuhl et al. define prototypes as “good instances of categories”, and these prototypes function as ‘perceptual magnets’ for other sounds of that category (1995: 123).

The NLM theory also aimed to account for the perceptual patterns of adult learners of a foreign or second language. This theory proposes that the acoustic correlates of the native language speech filter the perceived foreign or second language (Kuhl 2000). Under the effect of the learner’s native language prototypes, L2 learners are predicted to have difficulty discriminating an L2 sound that is phonetically close to a native language prototype. Learners are also predicted to easily discriminate an L2 sound that is phonetically different from any L1 prototype. For example, the difficulty that the Japanese learners of English have in perceiving and producing English /ɹ/ and /l/ was accounted for by the NLM theory as a result of Japanese language prototypes, which was suggested to have manipulated the English categories and this caused overlap in the perception and production of these categories by Japanese learners (Iverson and Kuhl 1995, Kuhl and Iverson 1995: 142).

Iverson et al. (2003) examined the perception of synthesised /ɹ/ and /l/ stimuli, which varied in F2 and F3 steps. German, Japanese, and American English participants were asked to identify the stimuli in terms of their L1 phonemes, and then they were asked to rate the goodness of fit of that category to their L1 sound. In a discrimination task, the participants were also asked to discriminate pairs of stimuli which differed in F3 only (half of the pairs were similar and the other half were different). The results showed that the Japanese listeners, who had one category to represent both /ɹ/ and /l/, were not sensitive to F3 differences. The Japanese participants were found to be more sensitive to F2 differences, which is irrelevant to the /ɹ/ and /l/ categorisation. Iverson et al. argue that the Japanese listeners filtered out the English categories through their perceptual space, in which they relied on irrelevant acoustic cues, F2, to form their own representations for /ɹ/ and /l/. Thus, the main insight of the NLM theory is that the native language prototypes act as magnets that filter the phonetic categories of the L2 (Kuhl and Iverson 1995: 144). This is particularly relevant for foreign language learners who live in a dominating L1 environment. Since the NLM theory was not aimed to L2 perception, it does not provide adequate predictions based on the context of L2 learning, in which factors other than L1 prototypes matter such as age of learning, L2 input and L2 use. So, the NLM does not make a suitable model to start with to account for FL speech. However, it can be referred to when it comes to L1 effect since its main focus is the L1 system.

2.2.2 Perceptual Assimilation Model (PAM)

This section presents background review of the Perceptual Assimilation Model (PAM) (Best 1994, 1995, 1999) which was used in this thesis to predict L2 perceptual patterns of English vowel contrasts by Syrian Arabic (SA) learners. Some of the predictions suggested by PAM were used in the methods chapters of this thesis.

The Perceptual Assimilation Model (PAM), which has been developed by Best (1995, 1999), was formulated to model foreign language speech perception which is referred to as cross-language speech perception. The principles of this model have been developed based on the principles of the Direct Perception proposed by Gibson (1966, 1979), for more details on Direct Perception see section 2.1.2 above (Best 1995). The main interest of PAM is speakers who have little experience with the foreign/or L2 language. It involves examining cross-language perceptual assimilation in a laboratory environment rather than in a natural L2 learning environment (Guion et al. 2000).

The PAM is predominantly a perceptual model which provides general predictions about the assimilation patterns of L2 learners. However, it does not provide any clear predictions about the production patterns of L2 learners. Regarding L1 acquisition, perceptual development often precedes production development (Best 1999). Following this sequential development of the L1, one can argue that L2 perception precedes L2 production. Thus, the production patterns of L2 learners can be inferred based on PAM perceptual predictions, which will be the base for the production task in this thesis.

Best claims that the perception of sounds in a foreign language is determined by their gestural similarities to, or discrepancies from, L1 sounds (Best 1994). The main claim of this model is that perceptual limitations of L2 learners determine the kind of difficulty the learners may encounter when learning L2 sounds. PAM is interesting, particularly, because it postulates that listening to L2 sounds does not only entail deciding which L2 sounds are similar to the L1 sounds and those which are different (Pilus 2005). For example, if two languages share a similar phoneme or phonemic contrast, it is likely that these languages differ in the articulatory and phonetic detail of those phonemes (Best 1994). Thus, listening to the L2 is a complex process that involves discriminating two L2 sounds from one another as well as distinguishing the L2 sounds from L1 sounds (Pilus 2005).

Best (1995: 193) defines non-native segments as ‘those whose gestural elements or intergestural phasing do not match precisely any native constellations’. Thus, the similarity between non-native and native segments is judged in the PAM based on their articulatory and gestural properties. Best (1995: 199) proposes that a non-native segment can be perceptually either (i) assimilated to a particular native segment as poor or good exemplar of that L1 category, (ii) assimilated as uncategorisable speech segment, (iii) or perceived as a non speech sound. According to PAM, a number of pairwise assimilation patterns have been suggested. Those patterns have been suggested based on the degree of similarity between L2 and L1 sounds, and the predicted discrimination of L2 sounds.

First, *Two-Category (TC)* members of the L2 contrast assimilate to two different native categories, that is, one member assimilates to one native category and the other one to another native category. The TC assimilation type yields excellent discrimination (Best 1995: 199). For example, Best et al. (2001) assessed native American English (AE) speakers' perception of three Zulu contrasts, in which voiceless vs. voiced lateral fricatives (/ɬ/-/ɮ/) were considered a TC pair. Zulu /ɬ/ consonant was predicted to map to AE phonological voiceless fricatives /θ s ʃ/ clustered with /l/, whereas /ɮ/ was predicted to map to AE voiced fricatives /ð z ʒ/ clustered with /l/. The principles of Articulatory Phonology (Browman and Goldstein 1989, Browman and Goldstein 1992) were used to choose the non-native contrasts. Based on the articulatory-phonetic characteristics of the contrasts under investigation, their similarity and discrepancy from their AE closest equivalents were determined, i.e. the articulators, constriction locations, and constriction degrees. The English listeners completed a categorical AXB discrimination test for the non-native contrasts, in which the participants were asked whether the target X is the same as A or B. The results revealed that the lateral fricative contrast was correctly discriminated 95% of the time. In order to evaluate the perceptual assimilation of these contrasts, the participants were asked to write what the non-native syllables sounded like to them in English orthography. The results coincided with the expected TC assimilation type for this contrast.

Second, *Category goodness (CG)*, in which each member of the L2 contrast assimilates to the same one native category with one of the members being more deviant from the native sound than the other. Thus, this category yields from moderate to very good discrimination (Best 1995: 199). For example, Zulu voiceless aspirated vs. ejective velar stops (/k^h/-/k'/) were suggested to form a CG pair, in which both were narrowly transcribed as [k^h] which is close to AE /k/ (Best, et al. 2001). The results of the categorical discrimination showed that AE participants correctly discriminated 89.4% of this contrast. Also, the perceptual assimilation patterns of this contrast were as predicted, in which both Zulu velars were assimilated to AE /k/ with a noticeable preference for the aspirated velar.

Third, *Single Category (SC)*, in which both L2 sounds assimilate to one phoneme in the native language and both are equally deviant from the native sound, and this yields poor discrimination (Best 1994, 1995: 199). Best et al. (2001) also examined the discrimination and perceptual assimilation of Zulu plosive vs. implosive bilabial stops (/b/-/ɓ/) which were analysed as a SC pair. The members of this contrast were both predicted to be equivalents to AE /b/. The results of the categorical discrimination showed that the participants did worse than the previous contrasts as 65.9% were correctly discriminated. The assimilation patterns of this contrast were as predicted for most of the participants as they were mapped into AE /b/ with no differences in the goodness of fit. However, other participants assimilated Zulu plosive /b/ with a closer goodness of fit into AE /b/, which suggests that this was a CG pair to them rather than SC.

Fourth, if the non-native sounds fall within the native phonetic space but phonemically uncategorisable to a particular native category, both sounds are analysed as *Uncategorisable* (UU) pair. The members of the UU pair are predicted to have from poor to very good discrimination. On the other hand, if one of the members of this contrast is phonemically categorisable into a particular native category, they will constitute *Uncategorised vs. Categorised* (UC) pair, which yields very good discrimination (Best 1995: 199).

Finally, if the non-native sounds are *Non-Assimilable (NA)* to any native category, they will be perceived with good discrimination as non-speech sounds (Best 1999). Best (1994) argues that adult listeners use the auditory differences in the discrimination of the members of the NA pairs. Best claims that since NA pairs are perceived as non-speech sounds, listeners will not use the phonetic and articulatory differences to discriminate the members of these pairs. An example of NA pairs is the Zulu apical vs. lateral clicks (/!/-/ǁ/), which was predicted to be NA to any AE category (Best 1994). The results of AXB categorical discrimination task showed that adult listeners correctly discriminated between 85-95% of these pairs. See Figure 2-1 below for an outline of PAM pairwise assimilation patterns and their predicted discrimination.

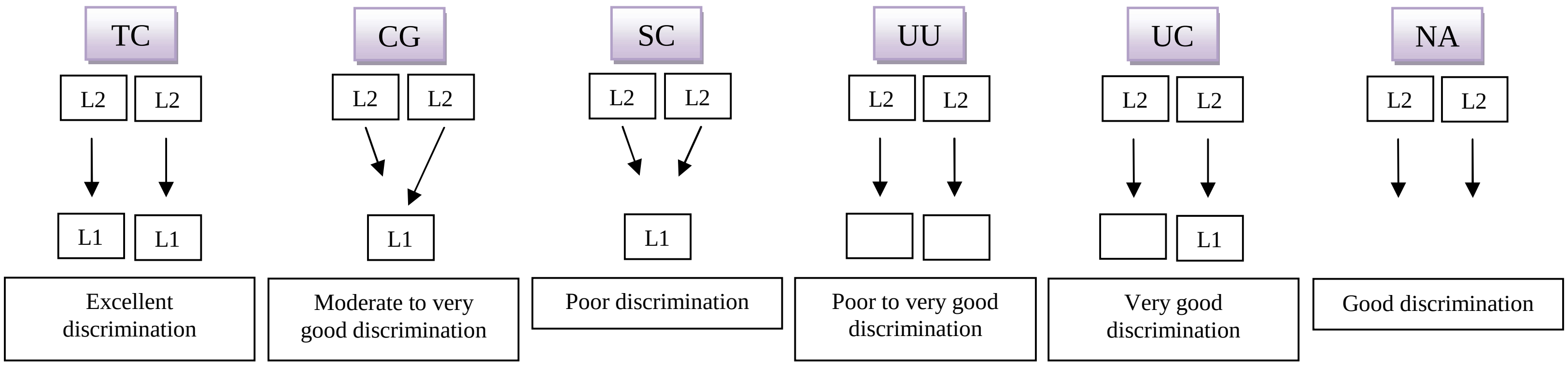

Figure 2-1: Outline of PAM pairwise assimilation patters and their predicted discrimination based on the similarities between L2 and L1 sounds

Best et al. (2001) predicts a gradient perception of TC>CG>SC contrasts, which is supported by the results of their experiment mentioned earlier, i.e. the Zulu lateral fricatives were discriminated better than the velar stops, which were discriminated better than the bilabial stops. They claimed that non-native listeners use phonological (TC pair which maps to a native phonological contrast), phonetic (CG pair phonetic goodness of fit), and non-linguistic information (SC pair in which listeners fail to detect the linguistic differences, so they refer to auditory differences similar to that of NA pairs) in their perception of the non-native sounds. They found near ceiling discrimination of non-native sounds if they were phonologically equivalent to their native sounds, lower but good discrimination if they were perceived with phonetic distinction CG, and poor discrimination if they were phonetically equivalent in goodness of fit to the listeners' native sounds. These findings suggest that experience with native language phonological as well as phonetic forms play a great role in the perception of non-native contrasts. Additionally, familiarity with the native language phonetic forms can be argued to have a positive effect on the discrimination of non-native contrasts. However, how positive this effect can be is subject to the phonetic similarities between the L1 and L2.

A number of studies have examined the perception of non-native contrasts based on the predictions of PAM (Almbark 2011, Best, et al. 2001, Ingram and Park 1997, Lengeris and Hazan.V 2007, Pilus 2005). For example, Pilus (2005) tested the identification and discrimination of voicing in three English pairs of word-final obstruents, (/t/-/d/), (/s/-/z/) and (/f/-/v/) by Malay learners of English. In Malay (/t/-/d/) contrast occurs in word-initial and medial position but not in word- final position. The (/f/-/v/) contrast does not exist in Malay because these consonants do not exist in the Malay phonemic inventory. The final contrast (/s/-/z/), does not exist either; the Malay phonemic inventory does not include /z/, and /s/ only exists in loan words from Arabic and English. Based on an identification and discrimination task, (/t/-/d/) and (/f/-/v/) pairs were classified as TC pairs, and (/s/-/z/) as a CG pair. In the identification task, the participants were asked to identify the word they heard from a minimal pair. In the discrimination task, the participants were presented with pairs of words and they were asked to decide whether they are the same or different. Accordingly, the analysed CG (/s/-/z/) pair was found to be the most difficult among the final consonant pairs, thus, supporting the PAM prediction. Since both L2 sounds /s/ and /z/ were assimilated to one native category, i.e. /s/, discriminating them is more difficult.

Furthermore, Ingram and Park (1997) confirmed the effect of L1 transfer and prior phonological learning which is the main interest of PAM. They examined cross-language vowel perception by adult Japanese and Korean learners of English. They were interested in examining the effect of prior phonological learning of temporal and quality features in the perception of Australian English vowels. Vowel length plays a great role in the phonological contrasts of Japanese vowel system, whereas Korean does not have phonological contrasts based on vowel length. Therefore, Ingram and Park (1997) argue that Japanese learners will be able to use this knowledge of vowel length to perceive non native phonological contrasts, whereas Koreans will not rely on vowel duration because it is not available to them in their L1. In terms of quality, the Japanese vowel system does not have an equivalent vowel to Australian English /æ/, whereas Korean has an (/e/-/ɛ/) phonemic contrast in which /ɛ/ was predicted to be an equivalent to Australian /æ/. In terms of PAM, Australian (/e/-/æ/) was argued to be a TC pair for Koreans and UC pair for Japanese. In both cases, the members of this contrast were predicted to be highly discriminable.

The participants were asked to identify the standard Australian English vowels by listening to vowel tokens and circling one of five choices on an answer sheet indicating which vowel they heard. Then, they were asked to give perceptual judgments of the same five vowels from the standpoint of the listeners' native language by transcribing the English item in Japanese Kana or Korean Hangul, and then rated how close each English vowel item matched the Japanese or Korean vowel chosen in their transcription. The results of their study show that L1 Japanese listeners outperformed the L1 Korean listeners regardless of English experience. The Australian (/e/-/æ/) contrast was found to present perceptual difficulties for Japanese and Korean participants. However, the Japanese listeners performed better than the Korean listeners in making this distinction although Japanese does not have a direct equivalent to Australian /æ/, and this does not coincide with PAM predictions. They attributed this result to the Japanese L1 "native-category transfer effects in the perception of non-native vowels (1997: 362)", i.e. the Japanese participants were predicted to have used durational differences between Australian (/e/-/æ/) vowels which was not available to Koreans.

Ingram and Park (1997) claimed that the durational cue present in L1 (Japanese) phonology led to a tendency toward categorical contrast in the perception of L2 (Australian English) vowel duration.

Best (1994) argues that adult learners of another language do not completely lose sensitivity to non-native contrasts. She claims that the effect of the experience with L1 phonological knowledge on the perception of non native sounds can be seen as an adjustment of selective attention to particular linguistic characteristics in the speech signal. Best also proposes that the lack of particular phonologically meaningful contrasts in a language does not necessarily mean that the speakers of that language are not exposed to the acoustic and phonetic characteristics of these/or some of these non-native contrasts. For example, the phonetic detail that is relevant to the phonemic contrast in one language might be relevant to allophonic rather than phonemic contrast in another language, which suggests that adult listeners of non-native contrasts are influenced not only by their native phonemic structure, but also by their native phonotactic, coarticulatory, and allophonic knowledge (Best and Tyler 2007: 17). The present thesis will present an evidence of allophonic L1 contrast that will have a role in the perception and production of a non-native phonemic contrast.

More recently, Best and Tyler (2007) proposed an L2 version of PAM (PAM-L2), which presents predictions on the perceptual patterns of L2 learners rather than monolinguals or foreign language learners with little L2 experience. The main interest of PAM-L2 is natural speech communication rather than laboratory artificial situations of the original PAM (2007: 18). In this paper, the authors highlighted some of the commonalities and complementarities with the Speech Learning Model (SLM), which will be presented in the following section. PAM-L2 focuses on two groups: late L2 learners and monolinguals, which differ in their L2 perception. The first group is argued to perceive L2 phonological categories, whereas the second group perceives L2 phonetic categories (2007: 14). Best and Tyler argued that these two groups differ in their L2 perception because monolinguals or naïve listeners do not realise the phonetic detail of the phonological categories of the L2. Therefore, naïve listeners were predicted to not differentiate between the phonetic and phonological levels of the L2 (2007: 23). On the other hand, late L2 learners were suggested to perceive L2 sounds at the phonological level based on their assimilation patterns to their phonologically equivalent L1 sounds.

However, the phonological assimilation between L2 and L1 sounds was not intended to entail necessarily phonetic similarity (2007: 26). Based on these principles, Best and Tyler set up PAM –L2 framework similar to the original PAM as follows:

1. TC or UC pair: only one L2 phonological category is assimilated as equivalent to a particular L1 phonological category. For these types, learners are predicted to perceive the L1 and L2 categories as phonologically and phonetically similar. Learners are predicted to discriminate the members of the L2 contrast (2007: 28-29).
2. CG pair: two L2 phonological categories perceived as similar to a single L1 phonological category, with one of the L2 categories being more deviant. The members of the L2 CG pair are predicted to be well discriminated but not as well as the TC pairs. Best and Tyler predict that a new phonetic and phonological category to be formed for the deviant L2 category, whereas no new category is predicted to be formed for the phonetically and phonologically closer to the L1 category (2007: 29).
3. SC pair: two L2 phonological categories perceived as equally good or bad equivalents of a single L1 phonological category. The members of the SC contrast are predicted to have poor discrimination (2007: 29).
4. NA pair: two L2 phonological categories which are not assimilated to any L1 phonological category. If each of these L2 sounds were assimilated to different set of L1 phonemes, they are easily discriminated and learned. However, if these sounds were assimilated to the same set of L1 phonemes, learners will have some difficulty discriminating them, as they are close to each other in the phonological space. Therefore, Best and Tyler suggest that a new single phonological category is likely to be formed to encompass both L2 sounds (2007: 30).

All in all, PAM makes a good model with its testable perceptual assimilation patterns and their predicted discrimination. However, determining the assimilation patterns and their predicted discrimination is greatly subjective and relative. In particular, analysing the results of the cross-assimilation, categorical discrimination, and identification tasks can be subjective and varies from one researcher to another, see for example Best et al. (2001), Pilus (2005), and Ingram et al. (1997). For example, PAM does not provide concrete numbers or percentages for the correct discrimination via which one can judge how good it is. If we take the TC pair, for instance, PAM predicts it will have excellent discrimination; what would be considered as excellent: 80%, 90%, or even 100% correct discrimination?

The research reported in this thesis uses PAM-L2 perceptual predictions and takes the perceptual results reported in this research a further step to predict L2 production patterns of the same English contrasts by SA learners. Performing this step will have critical implications on the discussion of the perception-production link, particularly that of adult L2 learners.

2.2.3 Speech Learning Model (SLM)

Another influential second language speech model is the Speech Learning Model (SLM), developed by Flege (1995). In addition to PAM, the SLM was utilised in the analysis of the results of L2 speech in this thesis, in particular, L2 productions. The SLM was developed to account for age-related limits on native-like production in L2 speech. The main focus of this model was bilinguals with a good amount of L2 experience rather than beginners (Flege and Mackay 2010). Flege argues that using L2 speech of beginners might reveal differences from native speakers that are due to learning in progress, not inability to learn. However, one can argue that even learning in progress of beginner L2 learners might show systematic patterns of perception and production of L2 speech. Also, after spending so many years in an L2 country, some L2 speakers continue to have a strong foreign accent, with production errors similar to those produced by beginner L2 learners.

Flege is particularly interested in the second language speech of immigrants, and those who acquired their L2 in a natural context rather than in a classroom (Flege and Mackay 2010). Flege claims that speech learning is a changing process. However, he emphasised the effect of Age of Arrival (AOA) to the L2 country on L2 pronunciation: “earlier is better (1995: 233)”. Additionally, SLM claims that accurate perception of L2 sounds is required to gain accurate production of those sounds. Thus, SLM focus is not only restricted to the domain of L2 speech perception, it also focuses on the domain of speech production, which can be inferred from the model postulates in Table 2-1 and hypotheses in Table 2-2 below.

Table 2-1: SLM postulates of second language acquisition (Flege, et al. 1995: 239)

	SLM Postulates
P1	The mechanisms used to acquire L1 sound continue over the life span, and can be applied to acquire L2 sounds.
P2	The phonetic categories of speech sounds are specified in long-term memory.
P3	L1 and L2 sounds are identified using phonetic categories established in childhood for L1 sounds.
P4	Bilinguals try to keep enough contrast between L1 and L2 phonetic categories, which exist in a common phonological space.

Table 2-1 above presents SLM four postulates which were proposed by Flege et al. (1995). All of these postulates are taken into consideration in the present thesis. As can be seen, the similarities and differences between L1 and L2 sounds are judged in SLM based on their phonetic and acoustic properties, which is unlike PAM and its articulatory gestures. Flege et al. argue in the fourth postulate P4 that bilinguals keep their L1 and L2 phonetic categories in a common phonological space, but try to keep L1 and L2 categories contrastive.

Table 2-2 below also presents SLM seven hypotheses listed in Flege et al. (1995). The seven hypotheses are relevant to the analysis of the results of the present thesis, except H4. Flege et al. claims in H4 that the ability to detect the phonetic differences between L1 and L2 sounds and between L2 sounds is dependent on experience with L2 sounds and Age of Learning (AOL), and in other studies on Age of Arrival (AOA) to the L2 speech community. They argue that as the AOL and AOA increase, the ability to detect the phonetic differences between L1 and L2 sounds decreases. Therefore, the essence of this hypothesis does not apply to the learners of this thesis, who had their L2 experience in their L1 country, in a classroom context.

Table 2-2: SLM hypotheses (Flege 1995: 239)

	SLM Hypotheses
H1	Perceptually L1 and L2 sounds are related at the allophonic level rather than at the phonemic level.
H2	For an L2 sound a new phonetic category can be established if it differs phonetically from the closest L1 sound, and if bilinguals could detect at least some of the phonetic differences between the L1 and L2 sounds.
H3	If the perceived phonetic dissimilarity between L1 and L2 sounds is great enough, L2 listeners will be able to detect the phonetic differences between the sounds.
H4	The ability to detect the phonetic differences between L1 and L2 sounds, and between L2 sounds that are non contrastive in the L1 decreases as AOL increases.
H5	Equivalence classification may block category formation for an L2 sound. In this case, L1 and 12 sounds will be perceived as diaphones[2], and they will be similar in production.
H6	The phonetic category established for L2 sounds by a bilingual may differ from a monolingual's. The bilingual tries to maintain phonetic contrast between categories in a common L1-L2 phonological space. The bilingual's representation might be based on feature weights.
H7	The production of a sound eventually corresponds to the properties represented in its phonetic category representation (which might not be native-like).

Since Flege et al. postulate that bilinguals have L1 and L2 phonetic categories in a common phonological space, they hypothesised in H1 that these categories are allophonically related. In H2 and H3 Flege et al. claim that L2 learners’ perception of L2 sounds is influenced by the phonetic categories of L1 sounds. During the early stages of L2 learning, L2 sounds are heard as instances of existing L1 categories regardless of the L1-L2 differences. After gaining some experience with L2 speech, adult learners may gradually detect the phonetic differences between some L2 and L1 sounds. At this stage, a new phonetic category representation might be established for the L2 sound.

For example, English liquids /ɹ/ and /l/ are heard by Japanese listeners as instances, i.e. allophones, of a single Japanese liquid /r/ (Flege, et al. 1995), with

[2] **‘Diaphone’ refers to two phonemes from different varieties or languages which are identified by a single phoneme in the native language or variety.**

English /l/ being perceptually more similar to Japanese /r/ than English /ɹ/ (Aoyama, et al. 2004). According to SLM H1 the English /ɹ/ and /l/ are perceived as allophones of the Japanese /r/. Since English /l/ is perceptually closer to the L1 sound, H2 suggests that Japanese learners of English should be able to detect some of the phonetic differences between Japanese /r/ and English dissimilar /ɹ/, and thus, form a new phonetic category to English /ɹ/ category. Using a categorical discrimination task, Japanese children and adult participants were asked to decide which one of three presented consonants is odd or not. This procedure was undertaken twice with one year separation between the two. The results showed that adult Japanese participants had no change in their discrimination of English /l/ and /ɹ/, whereas children's discrimination did improve, which was taken as a sign of improvement in their detecting of the phonetic differences between their L1 /r/ and English dissimilar /ɹ/ (Aoyama, et al. 2004).

Aoyama et al. (2004) also examined the productions of English /l/ and /ɹ/ consonants by the same Japanese participants. The productions were analysed twice (with one year time difference). Native English listeners were then asked to identify the initial consonants in the produced CV words. The results showed that Japanese children improved in their productions more than adults from time 1 to time 2, particularly for English /ɹ/ (35.1%). The categorical discrimination and production tasks supported the SLM hypothesis for the learnability of the dissimilar L2 sound. However, adult Japanese participants did not show a great improvement, which was claimed to be a result of more native input for the Japanese children than for the adults.

Flege et al. (1995) argue in H5 that establishing phonetic categories for L2 sounds might be blocked due to an equivalence classification with L1 sounds; that is L2 learners may fail to detect the phonetic differences between some L2 and L1 sounds, as they perceive L2 sounds as allophones of their L1. Thus, they perceive L2 sounds as instances of their equivalents in L1, which leads to lack of pronunciation accuracy and foreign accent in L2 speech (Bohn and Flege 1992, Flege 1987). For example, Flege (1987) showed that highly experienced English speakers learning French produced the French /y/ vowel, which does not have an equivalent in English, in a native-like way, whereas they produced French /u/ vowel, which is phonetically equivalent to English /u/, with an English phonetic properties. Therefore, Flege et al. argue that L2 learners with sufficient L2 input can establish phonetic categories for the new sounds if they are not in equivalence classification with any L1 sound.

Similarly, Bohn et al. (1992) hypothesised that the amount of L2 experience will have a great effect on the productions of English /æ/ vowel, which is a new sound to L1 German participants, whereas L2 experience will have no effect on similar English vowels such as /i/, /ɪ/, and /e/. In order to test this hypothesis, an acoustic analysis was conducted to compare the productions of these vowels between experienced and inexperienced German learners of English compared to those of native English. The results showed that both groups did not differ from each other in the acoustic properties of their productions of the similar sounds, whereas the experienced group produced the new English vowel /æ/ close to native English productions, which supports the model.

Additionally, an intelligibility labelling task was conducted to compare the intelligibility of the productions of the two German groups by three American listeners. The listeners were presented with /bVt/ tokens and were asked to choose the word they heard from six options. The intelligibility results did not show any significant differences between the two groups in the productions of /æ/ vowel, which does not support the model. Bohn et al. (1992) attributed this discrepancy between the two experiments, regarding the new category, to using different phonetic criteria to identify the productions of each group, or to perceiving consistent shorter durational L2 productions of /æ/ by both German groups, which led to similar intelligibility results.

Flege et al. highlight in H6 that bilinguals differ from monolinguals in their representations of sound categories. Monolinguals are argued to have learnt to attend to features (quantitative or qualitative) that distinguish contrastive sounds, and they tend to weight these features based on their role in the L1 phonetic system. Bilinguals attempt to keep L1 and L2 sounds contrastive. However, assuming that they keep the categories of their L1 and L2 in a common phonological space, the bilinguals are forced to have different features, to keep contrast L1-L2 sounds, than those used by monolinguals. Thus, it can be argued that having different features and feature weights for bilinguals entail having different deflected phonetic implementations in the productions of these L2 categories. Consequently, L2 speech might be perceived accented based on the role of those features in the perceived accentedness.

An advantage of the SLM is that it explicitly links L2 perception and production. Flege et al. claim in H7 that L2 categories will be ultimately produced with phonetic implementations that correspond to their feature or feature weight representations. However, based on H6 those features might not be native-like since their representations differ from the monolinguals'. This hypothesis is particularly relevant to the present thesis which investigates not only L2 vowel perception, but also L2 vowel production. Therefore, the ultimate aim of this thesis is to examine the perception-production link of L2 vowels. Based on SLM H7, L2 production patterns of the SA participants are predicted to *eventually* correspond to their perceptual patterns, i.e. the vowels that are easy to perceive will be easy to produce and vice versa.

Regarding vowels, which are the main interest in this thesis, the SLM assumes that adult learners may discern the phonetic differences between certain L1 and L2 vowels, and that the ability to discern these differences increases if the L1 has fewer vowels than the L2 (Flege 1995). Based on this assumption, Arab learners of English, which is rich in its vocalic system, particularly SSBE, are predicted to discern the differences between their L1 and their English L2 vowel systems since Arabic has a limited three-quality (or five-quality if it is an Arabic dialect) vowel system. This assumption will be clarified in the light of the present thesis.

With regard to the FL learners, the SLM cannot be applied directly as it builds most of its predictions based on AOL and AOA to the L2 country, as well as native L2 input. Since none of these factors apply to FL learners, the PAM will be mainly used in this thesis. Some reference can be made to the SLM as its learners share with FL learners the L2 knowledge, though not within the same context, i.e. SLM learners are naturalistic learners whereas FL learners are classroom learners.

2.2.4 A comparison between NLM, PAM and SLM

In the previous sections, a review of the main three L2 learning models was presented. These models, i.e. NLM, PAM, and SLM, were used in the analysis of the results of the perceptual and production experiments of the present thesis. Insights from the three models were used in the analysis of the results because the L2 learners of this study do not directly match the kind of learners that any of these models is interested in.

Table 2-3 below presents a comparison among the three models, which shows that PAM is particularly interested in naïve listeners, whereas SLM is interested in advanced L2 learners (Best and Tyler 2007, Flege 1995). On the other hand, the NLM theory is not developed to account for L2 perception; it is mainly interested in L1 perceptual development, which has specific implications for L2 perception (Kuhl and Iverson 1995).

Table 2-3: A comparison between the main three L2 speech perception and production models used in this thesis: NLM, PAM, and SLM

	L2 Model		
	NLM	PAM	SLM
The kind of L2 learners	L1 cognitive development (with implications for L2 learning)	L2 learners with little or no L2 experience	advanced L2 learners or bilinguals
Objects of speech	phonetic/acoustic properties of native language prototypes	actual articulatory gestures	phonetic/acoustic properties of speech signal
Level of perception	phonetic perception of L1 instances leading to phonological representations of prototypes	phonetic perception of L2 by naive listeners, and phonological perception of L2 by naturalistic L2 learners.	phonological perception of L2 by naturalistic L2 learners.
Perception-production link	perceptual model	perceptual model	links perception and production

Guion et al. (2000) examined whether SLM could be extended to inexperienced L2 learners. They investigated this possibility with three groups of Japanese speakers varying in their L2 experience. In a cross-language mapping experiment, near-monolingual Japanese listeners were asked to identify English (/b/, /v/, /w/, /θ/, /t/, /s/, /ɹ/, /l/) and Japanese (/b/, /ɰ/, /t/, /d/, /s/, /ɾ/, /h/) consonants in terms of a Japanese category followed by /a/ vowel (/ɰa/, /ɾa/, /ɯa/, /ɯɾa/, /ɾʲa/, /ja/, /ha/, /pa/, /ba/, /va/, /ɸa/, /ta/, /da/, /sa/, /za/, /ʃa/, /tsa/, /a/), and then, they rated the goodness of fit of their identifications. They predicted the consonants which had a good fit with L1 sounds to show little evidence of learning such as English /b/, /v/, and /s/, whereas the consonants with low fit with L1 sounds such as /ɹ/, /l/, and /θ/ were predicted to show evidence of learning.

A categorical discrimination test was used to examine the phonetic distance between English consonants and their closest equivalents in Japanese. They tested contrast pairs of four English-English consonants (/b/, /v/), (/ɹ/, /l/), (/s/, /θ/), and (/ɹ/, /w/), three Japanese-Japanese consonants (/ɾ/, /d/), (/s/, /h/, and (/s/, /d/), and seven English-Japanese consonants (/ɹ/, /ɾ/), (/ɹ/, /ɰ/), (/l/, /ɾ/), (/θ/, /s/), (/v/, /b/), (/t/, /t/), and (/b/, /b/). Based on the results of this study, they claimed that the perceived phonetic distance in the first experiment predicted the discrimination of L2 sounds in the second experiment. However, the perceived cross-language similarity of English and Japanese consonants did not predict the relative learnability of the consonant contrasts. Thus, their conclusion was that the SLM needs further investigation in order to be extended to early stages of L2 speech acquisition.

In the same study, Guion et al. (2000) examined whether PAM predictions could be extended to the natural L2 learning environment. The English (/ɹ/-/l/) contrast was predicted to have poor discrimination because these sounds occur between two Japanese categories (/ɯɾ/-/ɾ/), which would be considered as Uncategorised-Uncategorised UU in terms of PAM assimilation types. English (/ɹ/-/w/) and (/θ/-/s/) were analysed as PAM Uncategorised-Categorised UC assimilation type. The UC contrasts were predicted to show moderate to good discrimination since the second members of these contrasts were assimilated to L1 consonants /ɰ/ and /s/, respectively, whereas the first members were assimilated to two L1 categories (/ɯɾ/-/ɾ/) and (/s/-/ɸ/), respectively.

As predicted, Japanese listeners poorly discriminated the UU contrast. However, the results showed that the discrimination of the UC contrasts depended on the phonetic distance between L2 sounds and their equivalents in L1. For example, English uncategorised /θ/ was close in the phonological space to the categorised /s/ sound. Therefore, English (/θ/-/s/) contrast was poorly discriminated due to overlap in L1 classification. On the other hand, English (/ɹ/-/w/) contrast was highly discriminated because the uncategorised /ɹ/ was not close to the categorised one /w/. Thus, similar to SLM, the perceived phonetic distance from the first experiment was able to predict the discrimination of L2 sounds in the second experiment. The results of this study suggest that the PAM framework can be extended to L2 natural acquisition with a minor revision, i.e. they suggest to examine in UC cases how close the two L2 sounds in the phonological space from their equivalents in L1.

As can be seen, there is a gap in the literature that the above models do not account for, which is classroom foreign language learners. Classroom foreign language learners represent a different type of learning with various environments, inputs, and motivations. Despite those variations, foreign language learners with similar L1 backgrounds seem to have very similar pronunciation errors or difficulties. It can be argued that this similarity stems from the effect of their L1 which is completely dominating their social and interactional life. In this sense, foreign language learners are similar to the type of learners PAM is interested in, i.e. both types of learners have little or no native L2 input. Therefore, PAM was mainly used to predict the perceptual patterns of the learners of the present study.

Additionally, Table 2-3 shows that L2 theories of speech perception differed mainly in the nature of the underlying perceptual representations of L1 and L2 phonological categories (Strange and Shafer 2008). The NLM theory and SLM argue that perceivers form their categories from the phonetic and acoustic cues of the speech signal, whereas the PAM suggests that perceivers extract the actual articulatory gestures from the speech signal (Best and Tyler 2007: 24). Since the present thesis is not primarily focused on the cognitive aspect of L2 learning, it does not aim to prove which of these theories is accurate in terms of the objects of speech perception. A more general view of speech production shows that the articulatory gestures (suggested by PAM) constitute the actual input whereas the acoustic signal (suggested by NLM and SLM) constitutes the output of speech. Therefore, whenever I refer to the objects of speech in the present thesis, I refer to the articulatory gestures as input of speech and the acoustic signal as the output.

The level of speech perception can be argued to differ based on the L2 experience the learners have (Best and Tyler 2007: 16-17), accordingly, the three models differ in their assumptions of the level of speech perception the learners may exhibit. Table 2-3 above shows that the NLM theory assumes phonetic perception leading to the construction of prototypes in L1. Thus, L2 perception should be at the phonological level according to the NLM theory since those adult learners come with their L1 prototypes. As for the PAM, Best et al. (2007: 16-17) argue that naive listeners use their phonetic perception due to the lack of the L2 phonological knowledge or exposure, whereas advanced L2 learners use their phonological perception which they gained from their learning the L2 in a natural setting. The SLM, which accounts for the advanced L2 learners, arguably assumes phonological rather than phonetic perception of the L2. With this in mind, the FL learners, who learned their L2 to satisfy particular educational requirements, differ from both naive and advanced L2 learners. The FL can be argued to have different level of L2 perception due to the lack of native L2 input, thus they are similar to naive listeners in using their phonetic level of perception, as well as to their phonological and structural knowledge which they gained through direct teaching in the classroom. Thus, the FL learners can be argued to exhibit both levels of speech perception. However, the extent to which their perception can be native-like can be subject to the phonological instructions they get in class.

Generally, the main interest of the above speech learning models has been to describe the processes and mechanisms used to perceive L2 sounds. The relationship between the perception of L2 sounds and their production has been tackled to differing extents in these models. However, no clear predictions about the production of L2 sounds have been suggested or tested. Table 2-3 above shows that both NLM and PAM were developed as perceptual models whereas the SLM explicitly expressed the link of perception-production. Therefore, the SLM was mainly used to analyse and discuss the results of the production experiment of the present thesis and, thus, the perception-production link.

Based on the literature of L2 learning (reviewed throughout section 2.2), the perception of phonetic contrasts is usually examined experimentally using identification and discrimination tasks. The results of an identification task are usually calculated by counting the number of correct responses by L2 learners compared to native participants (Strange and Shafer 2008). As for the stimulus material, cross-language studies usually use synthetic stimuli in order to control some of the acoustic parameters. However, the synthetic stimuli used in most of these studies sound artificial. Additionally, the phonetic or acoustic properties used by native listeners to perceive real speech might differ from those controlled in the synthetic stimuli. Therefore, real speech which is extracted experimentally will be used in this thesis.

Another debatable issue related to stimulus material is the use of real vs. nonsense words. Arguably, real items reflect lexical knowledge as well as phonological, whereas nonsense items reflect the phonetic and phonological knowledge. For the purpose of FL learners' studies, their knowledge is based on what is taught to them in the classroom, which is basically lexical, and they lack native L2 input from which they gain detailed phonetic and phonological knowledge. Therefore, using real words is more representative of the participants' knowledge and thus is more reflective of their problems (Strange and Shafer 2008). Thus, an attempt will be made to use real words that are familiar to the participants. The style of the stimuli presented also has an effect on the perceptual studies. For example, the outcomes of read speech stimulus material are not generalisable to continuous speech. Nevertheless, the results of read speech can be generalised to the classroom situation, in which a more formal and standardised style of interaction is used (Strange and Shafer 2008).

The present study is an examination of the perception and production of English vowels by Syrian Arabic learners. The following sections provide a review of the literature related to Standard Arabic and SSBE vocalic systems. These two languages have been chosen for investigation because of the great difference between both systems in terms of quality and quantity of their vowels. Syrian Arabic dialect is examined in this study rather than Standard Arabic, since it has not been investigated in previous research and it is believed to have a slightly different vocalic system from the Standard Arabic. After that, a review of the available literature on L2 studies of Arab learners of English is presented.

2.3 Arabic Vowels

The rest of the background chapter will review the vowel system of the first and second languages under investigation, which are Arabic and English, respectively. A review of the quantitative and qualitative characteristics of Arabic vowels is discussed, in general, with particular focus on Syrian Arabic dialect. After that, a review of the quantitative and qualitative characteristics of English vowels with particular focus on Standard

Southern British English (SSBE) is presented in section 2.4, followed by a review of the L2 literature on Arab learners of English.

Arabic is one of the Semitic languages, which is spoken in twenty five countries in the Middle East and North Africa extending from the Arabian Gulf in the east to the Atlantic Ocean in the west. It is a widely spoken language, and learning it is important for Muslims as it is the language of the Holy Qur'an. The Classical Arabic form descended from the language of the western Hijazi tribe of Quraysh. The standard variety of Arabic, or Classical Arabic, is used in formal contexts such as education and broadcasting. Normative dialectal varieties, used for daily and informal communication, coexist with the standard variety, resulting in a phenomenon called **diglossia**, in which the two varieties (i.e. Classical Arabic and dialectal Arabic) coexist. The dialectal varieties of Arabic differ from standard Arabic in phonology, morphology, syntax, and lexicon. Recently, the term Modern Standard Arabic (MSA) has emerged; it refers to the Standard Arabic which shows the norms of the speaker's dialect (Watson 2002: 8). Standard Arabic is mutually intelligible among the speakers of all dialects. However, the dialects of extreme eastern and western Arab countries are mutually unintelligible. Additionally, Classical Arabic is never acquired as a mother tongue; rather it is taught formally at school. The mother tongue, in this case, is argued to be the dialectal variety of the speaker's region (Watson 2002: 8).

Newman & Verhoeven (2002) conducted an acoustic analysis of Arabic vowels to examine whether the prestigious high style of Classical Arabic or MSA could be used as an overall reference for the different varieties of dialectal Arabic. In their study, they were interested in vowels in connected speech rather than in laboratory speech. They used Quranic recitation style as it is considered the purest and most prestigious style. They analysed 30 minutes of recorded recitation, and they extracted 400 vocalic items representing the different vocalic qualities. Vowels in pharyngealised context were excluded from the analysis because of the dramatic effect of pharyngealisation on the frequency of these vowels. Additionally, an acoustic analysis was conducted for Egyptian Arabic, which was chosen for its prestigious position among other varieties of Arabic. The same vowels were extracted from connected speech for which speakers were asked to translate a text from English into Egyptian Arabic.

The results of the acoustical analysis of the Quranic recitation style vs. the vernacular style showed that Classical Arabic vowels were not at acoustic extremes. Furthermore, they found that the durational difference between long and short vowels was not significantly different. This finding contradicts the highly significant durational difference that was found in other studies of MSA or Arabic dialects (Allatif 2008, Almbark 2011, Khattab 2007, Khattab and Al-Tamimi 2008). This contradiction might be due to the use of isolated vowels or laboratory data in those studies that would, arguably, raise the durational difference between long and short vowels. Newman & Verhoeven (2002) claimed that the high style of Qur'anic recitation cannot be used as a reference to other varieties of Arabic. Thus, one can argue that examining the characteristics of Arabic vowels of the different varieties separately from Standard Arabic and from each other is required, and this is one of the aims of the present study.

In this section, a review of the literature of Arabic vowels is presented. The review covers previous research that was conducted on Classical Arabic and Dialectal Arabic. Vowel quality, vowel length, and diphthongs in Arabic will be reviewed. Finally, a detailed description of the Syrian Arabic vowel system, in particular, will be presented since it is the focus of the present research.

2.3.1 Vowel Quality

The vowel system of Classical Arabic or Modern Standard Arabic (MSA) is a simple three quality vowel system consisting of the high front /i(:)/, high back /u(:)/, and low /a(:)/ vowels with short/long distinctions (Mitchell 1993: 138, Newman and Verhoeven 2002). Thus, the fundamental short vowels /i, a, u/ and their long counterparts /iː, aː, uː/ constitute the Classical Arabic and the MSA vowel systems. In terms of quality, the Arabic vowel inventory is below the mean among UPSID languages[3], of which only 5.4% have a three-vowel system. As would be expected, the three vowel qualities of Arabic are also the most common three qualities (/i/, /u/, /a/) (Newman 2002). Some researchers argue that the short vowels do not have the same exact quality of their long counterparts. Watson (2002: 22) describes the /iː/ and /uː/ articulation as being closer

[3] UPSID (UCLA Phonological Segment Inventory Database) a database that contains the phoneme inventories of 451languages.

than that of their short cognates, and the /aː/ articulation as being fronter than its short cognate. Arabic vowels occur word-medially and finally. They are usually preceded by a glottal stop if they occur initially or after a pause (Kopczynski and Meliani 1993a).

As mentioned earlier, a great number of local varieties coexist with Standard Arabic. The spectral and temporal properties of the vowel system of these varieties vary due to linguistic and extra linguistic factors. For example, Arabic dialects do not all share the same exact vowel system that MSA has; some have one or two additional vowel qualities in the mid (/e/ and /o/) and central (/ə/) areas. For instance, the Jordanian Arabic vowel system is assumed to consist of /i iː eː a aː oː u uː/, Moroccan Arabic is assumed to consist of /iː aː ə u uː/ (Al-Tamimi 2007 a), and Syrian Arabic is assumed to consist of /i iː e eː a aː ə o oː u uː/ (Cowell 1964).

Arabic vowels usually have pharyngealised allophones, when they are adjacent to the pharyngealised consonants /tˤ dˤ ðˤ sˤ/. Pharyngealised vowels are more retracted than their non- pharyngealised cognates, and they are accompanied with a constriction in the upper pharyngeal area (Al-Ani and El-Dalee 1984, Jongman et al. 2007, Khattab et al. 2006, Kopczynski and Meliani 1993a). Generally, changing or adding a constriction in a different location results in more obvious auditory shifts of the second formant than in the first formant (Rosner and Pickering 1994: 27). Acoustically, pharyngealised vowels exhibit significant downward shifts of the onset of the F2 values and slight upward shifts of the onset of the F1 (Al Masri and Jongman 2003, Almbark 2008, Khattab, et al. 2006). In articulatory terms, moving the tongue from its retracted position into the position of the vowel takes time. For this reason, the quality and the quantity of the following vowel is argued to determine the degree of the effect of the pharyngealised consonant. For example, front vowels are assumed to be affected more than back vowels, low vowels are affected more than high vowels, and short vowels are affected more than their long counterparts. Thus, the low short vowel /a/ should show more evidence of pharyngealisation than any other vowel (Albashir 2008: 66-71, Gairdner 1925).

The symbol /ɑ/ is used by some researchers for the pharyngealised allophone of the plain low vowel /a/ (Gairdner 1925, Kopczynski and Meliani 1993a). However, the use of /ɑ/ symbol indicates that this vowel is more affected by neighboring pharyngealised consonants, and this effect is audible to native listeners i.e. native speakers are able to perceive the difference between /a/ and its pharyngealised allophone /ɑ/. Gairdner (1925) also describes the articulation of /a/ as being similar to the IPA /ʌ/ and the articulation of /aː/ as being similar to the IPA /æ/. However, further research is required to confirm Gairdner's descriptions.

The effect of pharyngealised consonants is not restricted to the following vowels. It is argued that pharyngealisation spreads in both directions. It can be anticipatory; however, pharyngealisation spread can be blocked in the context of high vowels /i/ and /u/, and even high consonants such as /j/ and /ʃ/ (Card 1983, Davis 1995, Watson 2002). In their studies on Palestinian Arabic, Card (1983: 49, 73) and Davis (1995) examined pharyngealisation spread by measuring the F2 of the following and preceding vowels. They found that pharyngealisation spreads in both directions rightward and leftward. However, they suggested that pharyngealisation spread can be blocked by high front vowels /y/ and /i/, and high consonants /ʃ/ and /j/. Additionally, Card found that F2 lowering of pharyngealisation is not required by /w/ and /uː/ since they are assumed to have low F2 values, but these segments do not block pharyngealisation spread.

In a study on pharyngealisation spread in two varieties of Syrian Arabic (Damascene and Aleppian), the productions of two male and two female speakers of each dialect were examined, using minimal pairs of the pharyngealised consonants and their plain counterparts (Almbark 2008). F2 lowering was used to test for pharyngealisation spread into the following and preceding vowels. The results showed that pharyngealisation spread in both directions. However, pharyngealisation was blocked by the high vowels /iː/ and /uː/. For /iː/, pharyngealisation was restricted to the following vowel only, but for /uː/ there was no significant difference from its plain counterpart. Thus, as in most studies of pharyngealisation, the effect of pharyngealised consonants is more evident on the low vowels.

2.3.2 Vowel Quantity

Generally, phonemic length has been reported to have a contrastive role in Arabic (Mitchell 1993: 138, Newman and Verhoeven 2002). The temporal opposition short vs. long applies to all consonants and vowels (Khattab 2007, Khattab and Al-Tamimi 2008). In terms of duration, Khattab (2007, Khattab and Al-Tamimi 2008) argued that long vowels are approximately twice as long as short vowels in Lebanese Arabic, a result that has been found for vowels occurring in spontaneous speech as well as in a word list task.

Phonemic length is affected by linguistic factors such as stress and focus, and voicing of the preceding and the following consonants. It is also affected by extra linguistic factors such as speech rate. The degree of stress has been reported to have an important role in the spectral and temporal properties of vowels (Flege et al. 1999). Jong and Zawaydeh (2002) examined how the interaction of stress and focus with segmental contrasts modifies vowel durations. In their study, minimal pairs were used to analyse the lexical and segmental focus effect in the speech of two male and two female Jordanian speakers in stressed and unstressed syllables. In the **lexical focus** frame, speakers focused on the target word, as in (1).

(1)	ʔana	ʔult	‘badawi’	miʃ	‘fallaħ’
	I	said	‘bedouin’	not	‘peasant’.

For the **segmental focus** the subjects corrected a listener who misheard a particular contrasting segment in the target word as in (2).

(2)	ʔana	ʔult	‘baada’	miʃ	‘bada’
	I	said	‘he extinguished’	not	‘he started’.

The results of their study showed that durational differences between long and short vowels are significantly larger in stressed syllables than in the unstressed syllables. They found that **lexical focus** is associated with increase in F1 values, and **segmental focus** rather than **lexical focus** increases durational differences. These findings coincided with what has been found for most languages (Jong 2004).

The effect of speech rate on vowel duration is also well documented in the literature (Allatif 2008, Rosner and Pickering 1994). For example, Allatif (2008) examined the correlates of vocalic quantity of three speakers from Mayadin, which is a city in eastern Syria. He tested the effects of speech rate and sentence type (declarative or interrogative). He used three minimal pairs for the extreme short vs. long vowel contrasts: /zir/-/ziːr/, /zar/-/zaːr/, and /zur/-/zuːr/, all target words were put in the carrier phrase /gaːl zVr/ ‘he said ___’. The results of speech rate showed that rapid speech reduced the duration of short vowels by 20% and long vowels by 19%. Additionally, all vowels were found to be longer in interrogatives than in declaratives. Generally, interrogation raised about 10% of the duration of all vowels. These findings are not surprising and meet our expectations, particularly for speech rate where vowels tend to shorten in rapid speech.

It is also universally accepted that the voicing of the following consonant has an effect on vowel duration. It has been reported that vowels are longer before voiced consonants than before voiceless ones (Denes 1955, Mitleb 1982, Munro 1993). Mitleb (1982) examined the productions of Arabic minimal pairs of eight Jordanian speakers. A spectrographic analysis revealed that voicing of the following segment did not have any effect on vowel duration. The effect of voicing on the preceding vowel was also examined in the productions of Saudi Arabic speakers by Flege (1979). The results of this study suggested that Arabic speakers do not exhibit a significant voicing effect on the preceding vowel. These results contradict voicing effect of the consonants on the preceding vowels which has been proved in other languages such as English. Generally, previous research on Arabic has suggested that Arabic is a quantitative language that relies extensively on the duration of segments to construct its phonological contrasts. It can be argued that due to the reliance of Arabic on duration to form phonemic contrasts, native speakers might aim to maintain the duration of vowels static and not changeable due to voicing of following consonants.

2.3.3 Diphthongs

Arabic has been reported in the literature to have no diphthongs as part of its vowel system. However, the two vowel-glide sequences /aj/ and /aw/ are sometimes analysed as diphthongs. These sounds were considered a special category of vowels in Arabic as they begin with one vowel and glide into a weak consonant or glide. This movement distinguishes these two sound sequences from the long vowels /iː/ and /uː/.

Some dialects of Arabic have diphthongs (/aj/ and /aw/) and/or monophthongs (/eː/ and /oː/) derived historically from the former diphthongs. The historical emergence of the two mid vowel qualities (/eː/ and /oː/) was assumed to be a result of a mutual interaction between two adjacent vowels or two components of a diphthong. For instance, in most Arabic dialects most /aj/ and /aw/ sequences in MSA evolved into /eː/ and /oː/, respectively. Table 2-4 below presents examples of MSA diphthongs and their monophthongal counterpart in Syrian Arabic dialect (Joseph and Odisho 2005: 49). This transformation is evident in many dialects in the Levant and in Cairene Arabic. However, the diphthongs are maintained in all phonological contexts in some of the peninsula dialects and in San’ani dialect (Watson 2002: 22).

Table 2-4: Examples of MSA /aj/ and /aw/ and their evolved form in Dialectal Arabic (Syrian Arabic) /eː/ and /oː/

Word	MSA /aj/	Dialectal Arabic /eː/	Word	MSA /aw/	Dialectal Arabic /oː/
'home'	bajt	beːt	'fear'	xawf	xoːf
'sword'	sajf	seːf	'sleeping'	nawm	noːm
'oil'	zajt	zeːt	'dress'	θawb	θoːb

There have been few attempts, if any, to analyse the distribution of the dialectal long mid vowels vs. the vowel-glide sequences of Standard Arabic. Gairdner (1925: 143) suggested that inherently formal lexical items adhere to the original vowel-glide sequences of Standard Arabic. The author gives /sawra/ 'revolution' and /qawmijja/ 'nationalism' as examples of such formal words. Nevertheless, this description is not satisfactory to illustrate the distribution of these sounds. Thus, it is needed to explore the phonetic and phonological features of diphthongs and their distributions in Arabic. Additionally, a phonetic and phonological analysis is required for each Arabic dialect vowel system to identify the qualitative and quantitative characteristics of each system including monophthongs and diphthongs.

Since the main interest of the present study is Syrian Arabic (SA), particularly the variety used in Damascus, the following section presents a brief description of the vowel system used in this dialect, which is mostly based on the auditory classifications of Cowell (1964).

2.3.4 Syrian Arabic Vowel System

In This section, a summary of the Syrian Arabic (SA) vowel system is presented based on the detailed auditory description of Cowell (1964). Cowell aimed to provide a detailed description of colloquial Arabic that can be used as a text book for L2 learners of colloquial Arabic. His descriptions referred to the conversational Arabic used in Greater Syria which includes Syria, Jordan, Lebanon, and Palestine. However, because of the great variability in colloquial Arabic of Greater Syria, his descriptions were based only on the variety used in Damascus the capital of Syria, which is located to the south as can be seen on the map of Syria in Figure 2-2 below.

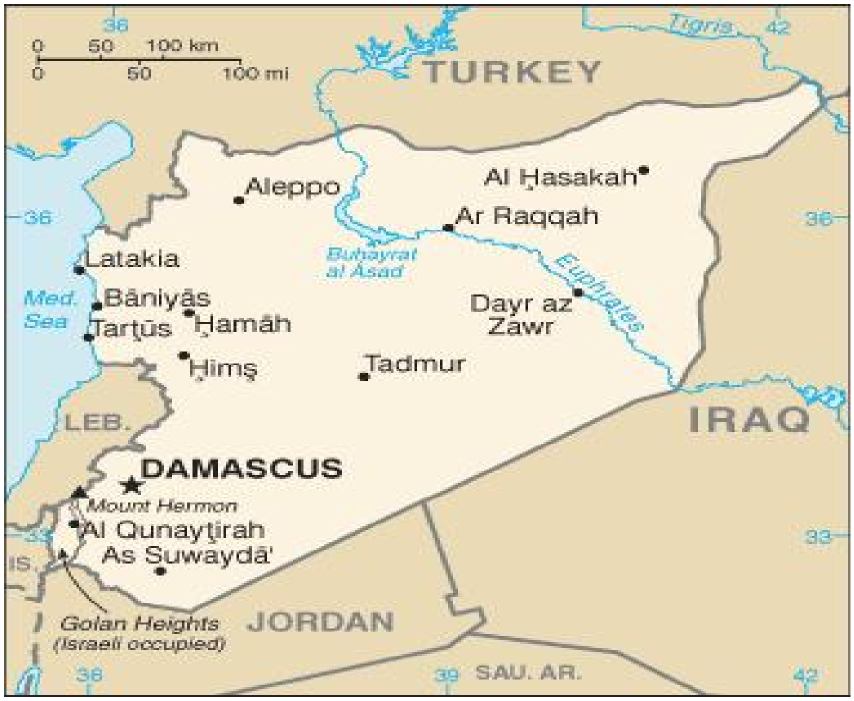

Figure 2-2: The map of Syria showing the capital city, whose dialect is the main focus of this study[4]

Based on auditory analysis of one male Damascene speaker, Cowell classified the speaker's vocalic productions into five long and five short vowels and a schwa. Table 2-5 below presents examples for each vowel, which are adapted from Cowell (1964). This description shows that the SA vowel system includes the fundamental short and long vowels of the MSA. Additionally, it includes the long mid vowels, which evolved from MSA diphthongs, and their short counterparts (see Section 2.3.3 above).

[4] This map was retrieved on 06-06-2010 from http://commons.wikimedia.org/wiki/File:Sy-map.png

Table 2-5: the set of vowels and some examples produced by the Damascene male speaker from Cowell (1964)

Short vowels	SA Examples	English Gloss	Long vowels	SA Examples	English Gloss
/i/	ʔiza	if	/iː/	kiːf	how
	sˤabi	boy		ʔiːd	hand
/e/	maːlek	owner	/eː/	beːt	house
	ħaːle	situation		zeːt	oil
/a/	ʕadd	counting	/aː/	faːz	he won
	kasuːl	lazy		ktaːf	shoulders
/o/	beːto	his house	/oː/	koːl	eat
	bjəmlok	he owns		xoːd	take
/u/	suʔaːl	question	/uː/	ʃuːf	look
	duwal	countries		bruːde	cold
/ə/	təlmiːz	student			
	sətt	lady			

Generally, the articulators during the production of the long sounds are argued to be held longer and tighter than during the short sounds (Cowell 1964: 15). For example, /iː/ and /uː/ are claimed to have more or less the same quality as /i/ and /u/, though the latter short vowels are believed to be slightly lower and tense. The short vowels /i/ and /u/ do not occur before a word final consonant. The long high vowels are close in their quality to English vowels. However, they are argued to be monophthongal in nature, i.e. they don't exhibit diphthongization which is an essential feature of English vowels (Cowell 1964, Gairdner 1925: 36). Being monophthongal or diphthongal is a feature which needs more investigation and acoustic analysis to examine the spectral properties throughout the vowel since neither Cowell and nor any other researchers examined this feature in Arabic to date.

The quality of the mid-high central vowel /ə/ is argued to depend, to a large extent, on the neighbouring sounds. Next to plain dentals /t d s n l/ or after /j/, /ə/ is believed to have a more forward value that is close to English /i/ in [sit], and lower value before pharyngeals /ħ ʕ/. As can be seen from Table 2-5 above, /ə/ is the only sound that has no long counterpart; and it never occurs word finally. Those descriptions of /ə/ together suggest that /ə/ has an underlying form, which can be one of the other short vowels. One can argue that, due to coarticulation effects, short vowels tend to centralise and surface in the form of /ə/ because short (lax) vowels are affected by the surrounding context more easily than long (tense) vowels (Rosner and Pickering 1994: 94). Gairdner (1925: 36) describes Arabic schwa as a vague vowel which can replace short vowels in rapid speech.

As for short mid vowels /e/ and /o/, they almost never occur accented and rarely in open syllables except word finally. The low short vowel /a/ is described as slightly raised and retracted English /a/. The long vowel /aː/ varies regionally as having higher and more forward values, such as /jwe̲ːmiʕ/ 'mosques', in Coastal regions than in Damascus.

All the above descriptions provide a general view of the vowel system used in Damascus. However, those descriptions are based on one male speaker only, whose productions might not be fully representative of his dialect, and this makes the descriptions not fully reliable. Obviously, the auditory analysis (or classification) was done by the author, however, the methods used in the recording, extracting, or classifying the data were not stated explicitly. Thus, the descriptions are subjective and require acoustical analysis to have an objective description of the vowels. Nevertheless, Cowell's description of Damascene sounds, in general, and vowels, in particular, can be used as a preliminary study that gives an adequate outline of the vowel system of this variety for further analyses.

SA vowels were also examined and analysed by Allatif and Abry (2004) and Allatif (2008). The productions of three speakers of SA from Mayadin were analysed. The three extreme short vowels of SA /i a u/ and their long counterparts /iː aː uː/ were examined in terms of their quantity and their quality. Three minimal pairs were used to extract the three short vs. long vowel contrasts, which are presented in Table 2-6 below.

Table 2-6: SA minimal pairs representing short vs. long vowels which were used in the study of Allatif and Abry (2004)

SA Short vowel	English Gloss	SA Long vowel	English Gloss
/zir/	button	/ziːr/	proper name
/zar/	he buttoned	/zaːr/	he visited
/zur/	you button!	/zuːr/	you visit!

All the target words were put in the carrier phrase /gaːl zVr/ ‘he said ___’. Their aim was to examine whether vocalic quantity is the only distinction between long and short vowels. The durations and the values of the first three formants of the vowels were extracted and compared. According to the results of this study, Allatif and Abry (2004) argued that both quantity and quality distinguish /a aː/ and /u uː/, with a greater role for quantity in the /a aː/ distinction and a greater role for quality in the /u uː/ distinction. The greater quantity distinction for the low vowels can be attributed to the fact that low vowels tend to be longer than high vowels, in general. Unexpectedly, only quality was claimed to have a role in the /i iː/ distinction. Relying on these results, their conclusion was that a contrast shift from quantity to quality is taking place.

Additionally, Allatif (2008) studied the effect of speech rate and structure on the fundamental frequency (F0) and intensity of the above vowels. He found that rapid speech and interrogation raised F0 and intensity significantly more than normal speech and declaration, respectively. Regarding the effect of speech rate and structure on the spectral cues, he found a robust F1 difference between short vs. long vowels, and robust F2 difference for /i/ only. All in all, Allatif’s (2008) results seems to support his (2004) results in terms of quality and quantity distinctions between short vs. long vowels.

The second part of Allatif’s (2008) study involved manipulation of the temporal cues of the vowels; short vowels were given the duration of long vowels and vice versa. For this study, he used 10 speakers from Mayadin × 20 words (10 manipulated and 10 controls) × 6 vowels. He used the same /zVr/ context as in the previous study; he measured the duration of prevocalic and postvocalic segments duration. The results showed that the duration of adjacent segments does not depend on the duration of the vowel. Thus, the duration of adjacent segments does not give any indication about the vowel short vs. long category. This result coincides with results of the previous research that voicing of adjacent consonants does not affect the duration of the following vowels in Arabic (Flege 1979, Mitleb 1982). He also found that F0 is independent from short vs. long categorisation, whereas intensity depends on vowel category only for (/i/-/iː/) distinction.

In sum, Allatif's work (Allatif 2008, Allatif and Abry 2004) suggests that the acoustic correlates of vocalic quantity for (/i/-/iː/) are (i) spectral, (ii) temporal, and (iii) prosodic (intensity), whereas for (/a/-/aː/) and (/u/-/uː/), the acoustic correlates of vocalic quantity are i) spectral and (ii) temporal. According to these findings, it is important not to ignore the quality difference between short vs. long Arabic vowels. The main focus of Allatif's research was the main three short and long vowels only though Syrian Arabic dialect is believed to have short and long mid vowels and a schwa (Cowell 1964). He also did not examine the vowels within pharyngealised context, which has a significant effect on vowels (see section 2.3.1). It is important not to ignore the mid vowel qualities /e o/ and pharyngealised vowels, particularly in L2 studies as they are predicted to aid in perceiving and producing some L2 vowels with similar qualitative and quantitative characteristics (Almbark 2011). Therefore, the present study aims to fill this gap by covering these vowels and examining the effect of pharyngealisation in perceiving and producing some L2 English vowels by SA learners.

Abou Haidar (1994) compared the productions of CVC words of the fundamental short vs. long vowels of MSA as produced by one speaker from different dialects (Qatari, Lebanese, Saudi, Tunisian, Syrian, Sudanese, Emirati and Jordanian). Figure 2-3 below presents the frequency values for the SA speaker from Abou Haidar's study at the top compared to the values of Allatif (2008) at the bottom. Similar to other studies, these studies used a limited number of speakers to represent the dialect. Figure 2-3 shows that Allatif's values (with a greater number of speakers) are more centralised than Abou Haidar's, particularly for short vowels. Nevertheless, these values cannot be generalised due to the limited number of participants in each study. Thus, a study with a considerably greater number of participants is required to generalise the acoustic characteristic of the vowels of SA, and this is one of the aims of the present study (see chapter 4).

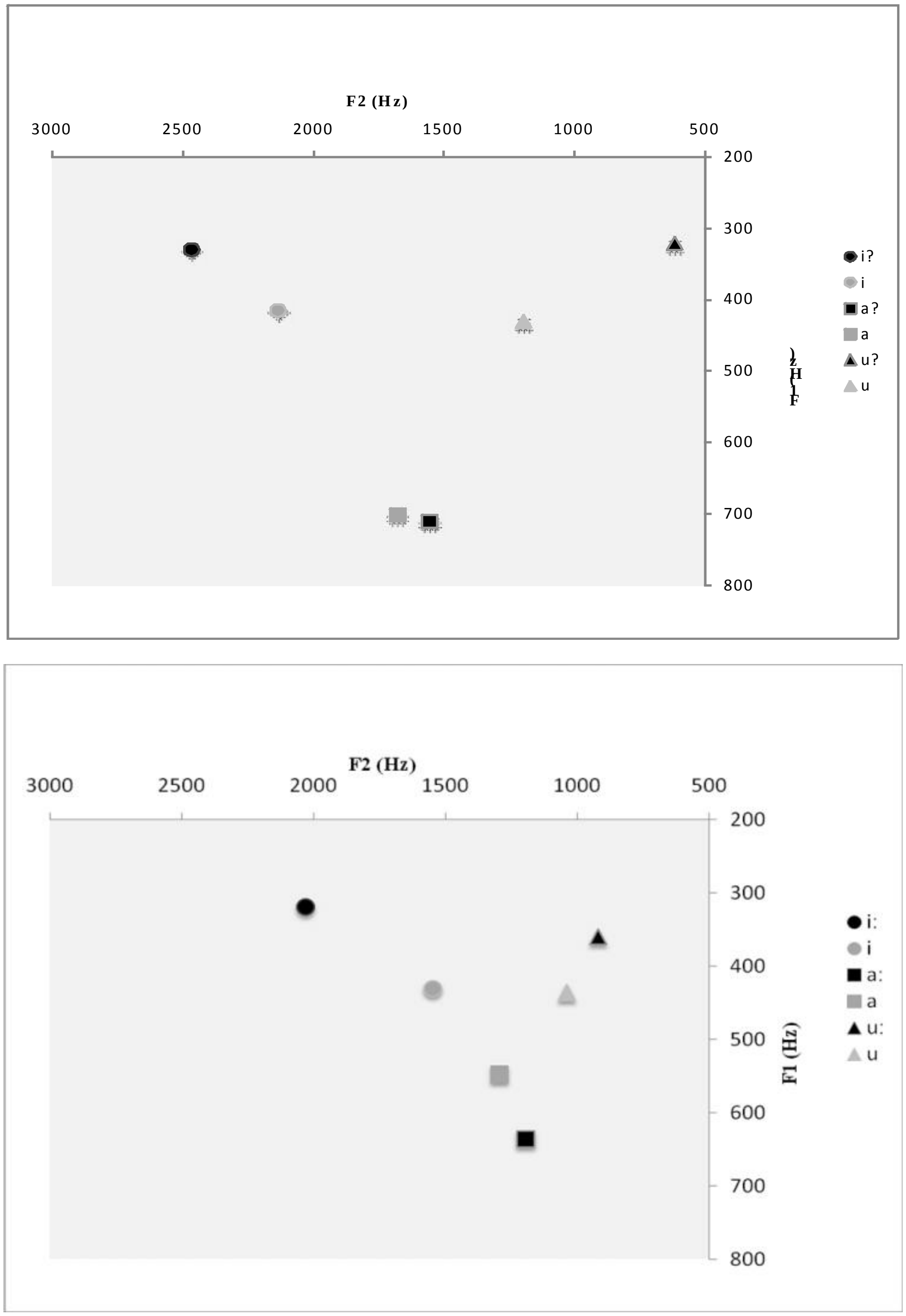

Figure 2-3: Mean frequency values of the basic short v. long vowels of Syrian Arabic. At the top the chart is based on one speaker from (Abou Haidar 1994), at the bottom the chart is based on three speakers taken from (Allatif 2008)

Following the review of the first language (L1) of the participants of the present study, i.e. SA, the following section presents a review of the vowel system of the second language (L2) of the participants under investigation, i.e. the Standard Southern British English (SSBE), including SSBE monophthongs and diphthongs.

2.4 English Vowels

Compared to MSA or any Arabic dialect, English has a more crowded vowel space, particularly back vowels. American English and British English are the most common English targets used in the education system of the majority of the Arab world, in particular, and in the education system of other countries, where English is taught as a foreign language. Standard Southern British English SSBE is claimed to be the target English language (L2) used in the Syrian Education system. For the last 6 years, a national communicative curriculum is compulsory from year one throughout the whole stages of education. It is called *English for Starters 6-9*, developed by Patricia Mugglestone in collaboration with York Press in the UK (Scobbie 2007). Therefore, the target L2 for the present study will be British English as it is the target language for English teaching in Syria. Due to the great variation between the vowels of the different English accents (Ladefoged 2006: 83), the focus of the present study will be the variety spoken in the south of England, London and its surroundings, i.e. SSBE.

However, the great influence of American English on the media led to an inevitable effect on L2 or foreign language learners, even though their target L2 is assumed to be another English variety. Arguably, this situation applies to Syrian Arabic learners of English. Therefore, caution will be taken into consideration during designing and analysing the L2 data of the present research.

The rest of this section presents a review of the literature examining and describing the qualitative and quantitative characteristics of SSBE vowel system. Additionally, the diphthongs of SSBE will be explored. Throughout these sections, the similarities and differences between SSBE and Arabic (in particular, Syrian Arabic (SA)) vowels will be highlighted and compared. These comparisons will set out the predicted perceptual and production patterns for the SA learners of the present study.

2.4.1 Vowel Quality

Wells (1986) presented a summary of the consonantal and vocalic phonological variation of several English accents, among which is SSBE. Wells set a list of standard keywords to represent each vowel, for a list of SSBE vowels and their keywords, which will be used throughout this thesis, see Table 2-7 below. Unlike Arabic, the English vowel system is claimed to be an example of a quality language; in which each vowel is produced with a distinct set of formants corresponding to its tongue height and advancement. Being a qualitative language, English is predicted to impose more difficulties on Arab learners who come from a predominantly quantitative language (Khattab 2007, Khattab and Al-Tamimi 2008, Mitchell 1993, Munro 1993).

Table 2-7: The vowels of SSBE and the standard lexical set adapted from Wells (1986: 123)

Tense vowels	Keywords	Lax vowels	Keywords	Diphthongs	Keywords
iː	FLEECE	ɪ	KIT	eɪ	FACE
ɜː	NURSE	e	DRESS	aɪ	PRICE
ɑː	BATH	æ	TRAP	ɔɪ	CHOICE
ɔː	THOUGHT	ɒ	LOT	aʊ	MOUTH
uː	GOOSE	ʌ	STRUT	əʊ	GOAT
		ʊ	FOOT	ɪə	NEAR
		ə	COMMA	ɛə	SQUARE
				ʊə	CURE

SSBE has been found to have a more crowded vowel space than other varieties of English. McMahon (2002: 96) argues that this crowded vowel space of SSBE is due to the fact that it is a non rhotic accent; therefore, it has all the centring diphthongs. She points out that as a result of rhoticity absence from the SSBE system, there are oppositions which exist in SSBE (e.g. TRAP-PALM) and not in other accents such as Scottish English. Lack of rhoticity in SSBE is interesting, particularly, within L2 learning when English is the target language for learners with a rhotic L1 background. For example, Arabic which is the main interest of the present study, is a rhotic language, and the question would be how do Arabic learners of English produce or perceive these English contrasts.

Another spectral property of English vowels is that English monophthongal vowels are claimed to exhibit no static state throughout the vowel. They tend to have spectral changes represented in formant movement. Andruski, et al. (1992) define what they call ‘vowel-inherent spectral change’ as “the relatively slowly varying changes in formant frequencies associated with vowels themselves even in the absence of consonantal context (Andruski and Nearey 1992: 393)”. However, diphthongisation is argued to be dialect dependent (Escuderoa and Chládková 2010). Escuderoa et al. (2010) argue that American English monophthongs exhibit formant movement more than SSBE.

On the other hand, Chladkova and Hamann (2011) predicted that for a case of near-merger in the F2 values of SSBE FLEECE and GOOSE vowels, resulting from GOOSE fronting, a possible acoustic cue to distinguish the two vowels would be the amount and direction of formant movement. In order to test their prediction, they have extracted F1 and F2 values at 25%, 50%, and 75% of the vowel duration. The results showed that the steady state (50%) can be still used as an acoustic cue to distinguish both vowels. Additionally, they found another cue to distinguish these vowels, i.e. the F2 contour which is falling for GOOSE and rising for FLEECE. Thus, not only diphthongs but also English monophthongs are suggested to have diphthongisation as a distinctive feature, which provides acoustic cues for the perception of English vowels. It can be argued that the lack of diphthongisation in the L2 speech of English learners could account, in addition to other factors, for the perceived foreign accent.

In English, there is a tendency to nasalise vowels before nasal consonants as in ‘beans’ and ‘song’. What happens is that when a vowel occurs between two nasals, before or after a nasal consonant, the velum lowers earlier during vowel production which allows air to release through the nasal cavity. This feature of English vowels is claimed to be allophonic and has no role in vowel identification. Similar to diphthongisation, the lack of nasalisation in the L2 speech could also have an effect on the perceived accentedness. L2 speech lacking such English features or phonological processes can be a natural product of learning a second language as an adult. Because, adults are not predicted to have access to the L2 phonological and allophonic processes;

all they have access to is their L1 phonemic, allophonic, and phonotactic processes (Best and Tyler 2007).

Deterding (1997) measured the frequencies of the eleven SSBE monophthongal vowels for five male and five female BBC broadcasters whose speech was taken from the MARSEC database. An average of ten tokens of each vowel for each speaker was extracted from connected speech. Figure 2-4 and Figure 2-5 below present the mean frequency values for the male and female speakers, respectively, which were extracted from connected speech. The aim of this study was to compare the measurements of the vowels in connected speech to the previous studies which focused on word citations. It also provided a reference for other researchers who wanted to examine SSBE monophthongs. As would be expected, the vowel space of females is more spread and peripheral than males'.

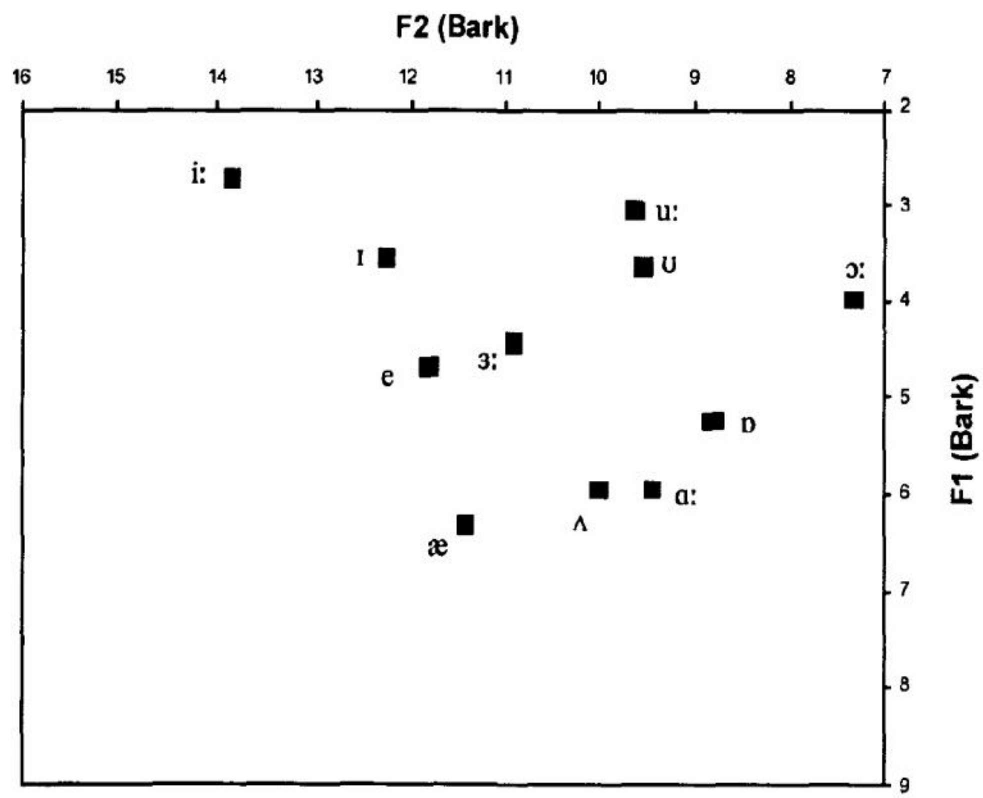

Figure 2-4: SSBE mean frequency values for 5 male speakers, adapted from Deterding (1997)[5]

---

[5] /e/ and /æ/ are sometimes transcribed as /ɛ/ and /a/, respectively.

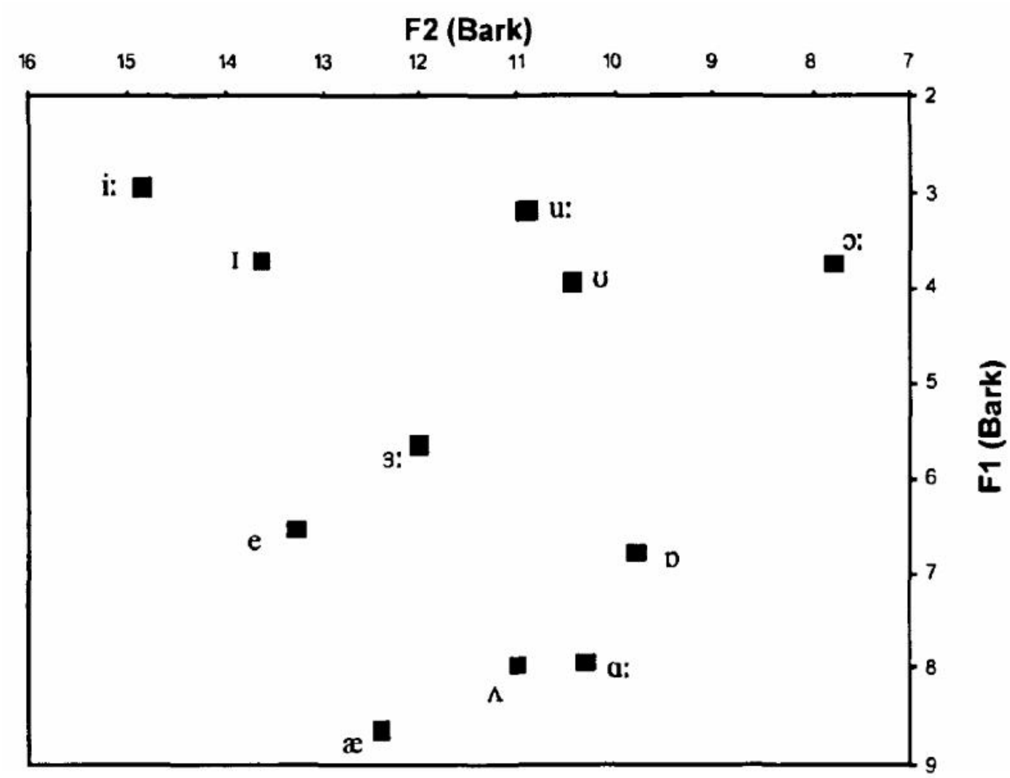

Figure 2-5: SSBE mean frequency values for 5 female speakers, adapted from Deterding (1997)

Another study examining the acoustic properties of SSBE monophthongs was conducted by Ferragne and Pellegrino (2010), in which they examined the vowels produced by males in 13 accents of the British Isles. Ten male and ten female participants from each region were recorded, including SSBE. The stimulus material was a list of /hVd/ words, among which nonsense words were used, but the participants were instructed to pronounce them similar to real rhyming words (e.g. hoid-void). Figure 2-6 below presents the F1-F2 values for the male participants. In terms of quality, the SSBE males' vowel space in Figure 2-4 (Deterding 1997) is comparable to that in Figure 2-6 (2010) though the numbers are not comparable. The above two acoustic studies of SSBE vowels can be used as a reference point to compare with the acoustic results of the present thesis.

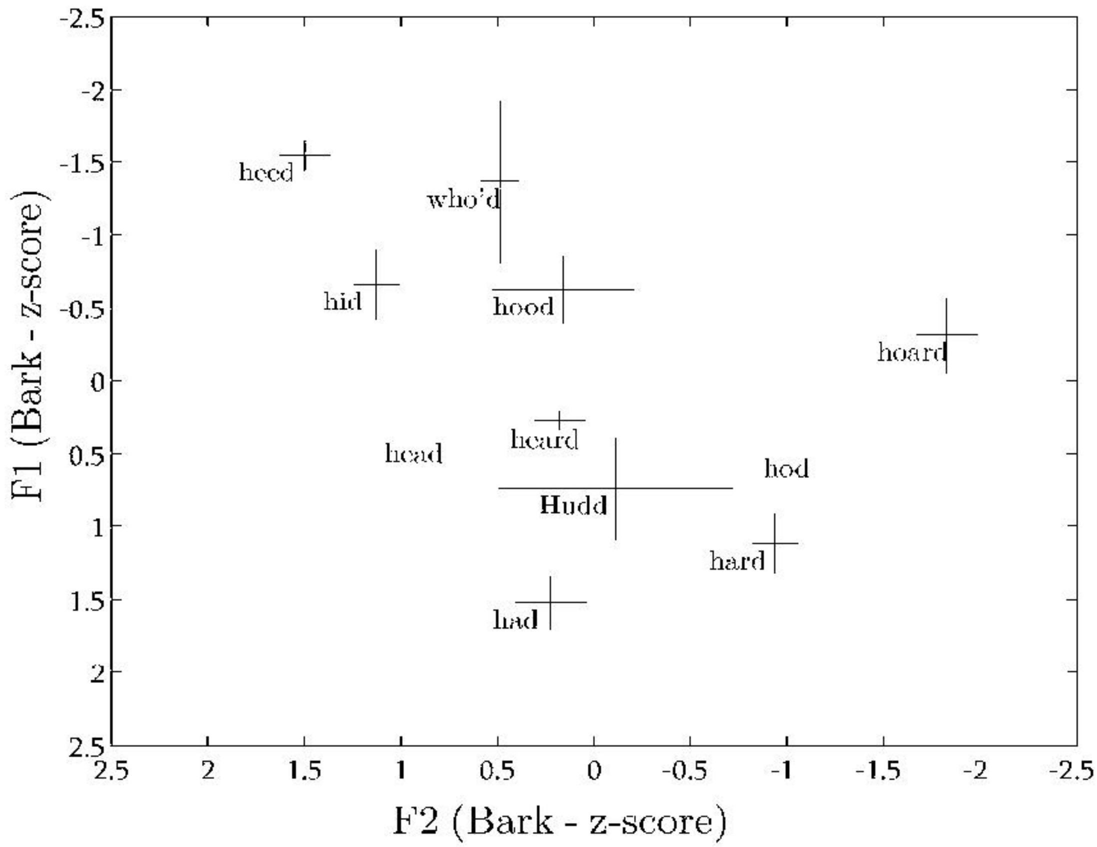

Figure 2-6: Median and interquartile range for the 11 monophthongs produced by six male speakers of SSBE, adapted from Ferragne and Pellegrino (2010)

Both male charts reflect some of the characteristics/changes of the SSBE monophthongs described by Hawkins and Midgley (2005), including what they argue to be well established changes in the speech of young people:

- Lowering of TRAP vowel (represented by /æ/ and 'had', respectively) into [a] vowel.
- Lowering of DRESS vowel (represented by /e/ and 'head', respectively).
- The use of a relatively central GOOSE vowel [ʉː] (represented by /uː/ and 'who'd', respectively)
- The use of a fronted and centralised FOOT vowel (represented by /ʊ/ and 'hood', respectively).

Other changes such as unrounding both GOOSE and FOOT vowels is claimed to be a recent innovation of the young people. Additionally, they referred to FLEECE, LOT and THOUGHT as the most stable SSBE vowels. Furthermore, Hawkins et al. (2005) descriptions of the changes of the SSBE vowels were supported by Cruttenden (2008).

Interestingly, such changes or innovations of the young speakers of SSBE are not predicted to be highly accessible to all kinds of L2 learners. For example, the SA learners of the present thesis do not have sufficient native L2 input. Thus, they are unlikely to perceive and produce the recent changes to SSBE vowel system. Moreover, English dictionaries which are designed for L2 and non-native learners use simplified phonetic transcriptions, which can be called canonical pronunciations. For instance, the GOOSE vowel is transcribed as /uː/ and not as its fronted form [ʉː], which can be argued to be sufficient for the L2 learning goal of those learners (Ferragne and Pellegrino 2010).

2.4.2 Tense vs. Lax Vowels

English vowels are associated with distinctive spectral and temporal properties. Some pairs of English vowels are distinguished by the feature of 'tenseness', which is associated with a more peripheral tongue placement and a slightly longer duration, such as /iː ɪ/ and /uː ʊ/, for a list of English vowels which are distinguished in terms of tenseness see Table 2-7 above. Chomsky and Halle (1968: 324-325) argue that the feature "tenseness" is a description of the articulatory gesture during the production of a given sound "Tense sounds are produced with a deliberate, accurate, and maximally distinct gesture that involves considerable muscular effort; no tense sounds are produced rapidly and somewhat indistinctly (1968: 324)". The suggested greater articulatory effort in the production of tense vowels is claimed to manifest in their greater distinctiveness and longer duration. The production of tense sounds requires the articulatory organs to maintain the appropriate configuration for a longer period of time than for nontense sounds.

Mitleb (1984b) claims that the durational difference between tense vs. lax English vowels is due to configurational obligation in which the articulatory organs maintain the appropriate configuration for tense vowels longer than for lax vowels; and this is different from the long vs. short feature of other languages such as Spanish, Arabic, and Japanese. Since English has been considered as a qualitative language, the phonological analysis of tense vs. lax contrasts has been mainly considered as a redundant phonetic difference that has no effect on the phonemic distinction of these vowels. However, this quantitative difference is predicted to be phonemically relevant to L2 learners with a quantitative L1 background, such as Arabic. Therefore, the SA participants of the present study are predicted to use their L1 long vs. short distinctions to perceive and produce their L2 tense vs. lax distinctions (Mitleb 1981).

Voicing of following consonants has been argued to have a significant effect on the duration of English vowels. Vowels which are followed by voiced consonants tend to be longer than vowels followed by voiceless consonants. However, these temporal differences related to voicing are allophonic and they don't interfere in the vowel phonemic identification (Arthur and House 1961). Therefore, the SA participants of the present study are not predicted to show an effect of voicing on vowel duration (Mitleb 1982).

2.4.3 Diphthongs

The SSBE vocalic system is rich in its diphthongal vowels due to lack of rhoticity. Table 2-7 above shows that English basically has 8 diphthongs. Diphthongs can be described with the trajectory direction they undertake, for example (FACE, PRICE, CHOICE, and MOUTH) are rising diphthongs, whereas (NEAR, SQUARE, and CURE) are centralised ones. Similarly, due to the lack of rhoticity, the English diphthongs PRICE, CHOICE, and MOUTH may occur followed by a schwa to compose triphthongs such as 'fire' /faɪə/, 'lawyer' /lɔɪə/, and 'power /paʊə/. Other researchers may analyse them as heterosyllabic; that is the diphthong and the schwa belong to two syllables (Chomsky and Halle 1968: 120, Wells 1986).

Deciding the perceptual cues of diphthongs has been a debatable issue. Gay (1970), using synthetic speech, found that the perceptual cue of diphthong identification in American English is the duration of the transition rather than the frequency of the onset or offset positions. On the other hand, Nabelek (1992) found that the correct identification of diphthongs lies in the onset steady-state and transition period since the offset segment steady-state was variable and not found in some tokens. This issue is particularly relevant to the present thesis, because focusing on a particular perceptual cue for diphthongs would have an effect on the analysis procedures of the perceptual and production diphthongal data, and thus, examine whether learners manage to use the same cues used by native speakers.

Most studies of English vowels focus on monophthongs. Few studies have been conducted to examine English diphthongs, particularly in SSBE. Diphthongal vowel contrasts have been investigated in other varieties of English such as American English (Gay 1968, 1970, Kiefte and Kluender 2005, Nabelek, et al. 1992, Peterson 1961). Even in L2 studies, the focus has been on monopthongs rather than diphthongs. This might be due to the complexity of analysing diphthongs. Thus, the present thesis can fill this gap, by examining the acoustic characteristics of SSBE diphthongs and their L2 production and perception.

In addition to SSBE monophthongs, Ferragne and Pellegrino (2010) examined the production of five closing diphthongs (FACE, PRICE, CHOICE, GOAT, and MOUTH). Figure 2-7 below presents the median of the starting and ending point of these diphthongs. As can be seen the trajectory movement for these vowels is clear, and learners of English are predicted to exhibit this movement in order to be perceived as native-like. As for the other diphthongs, the SQUARE vowel was argued to appear mostly as a monophthong that would be transcribed as [ɛː], which contradicts the traditional transcription of SQUARE as a centring diphthong /ɛə/ (Wells 1986). As for the NEAR diphthong, it appeared mostly as a centring diphthong. However, in some cases it was argued to be monophthongal. The CURE vowel was found to be problematic and exhibited a lot of variation, which was related to the use of the nonsense word ‘hured’ to represent this diphthong, which made it hard for the participants to decide the phonological status of this vowel. Hawkins and Midgley (2005), reported that one of the ongoing changes in the speech of young SSBE speakers is the loss of the schwa off-glide in CURE words, which is argued to surface now as THOUGHT vowel instead.

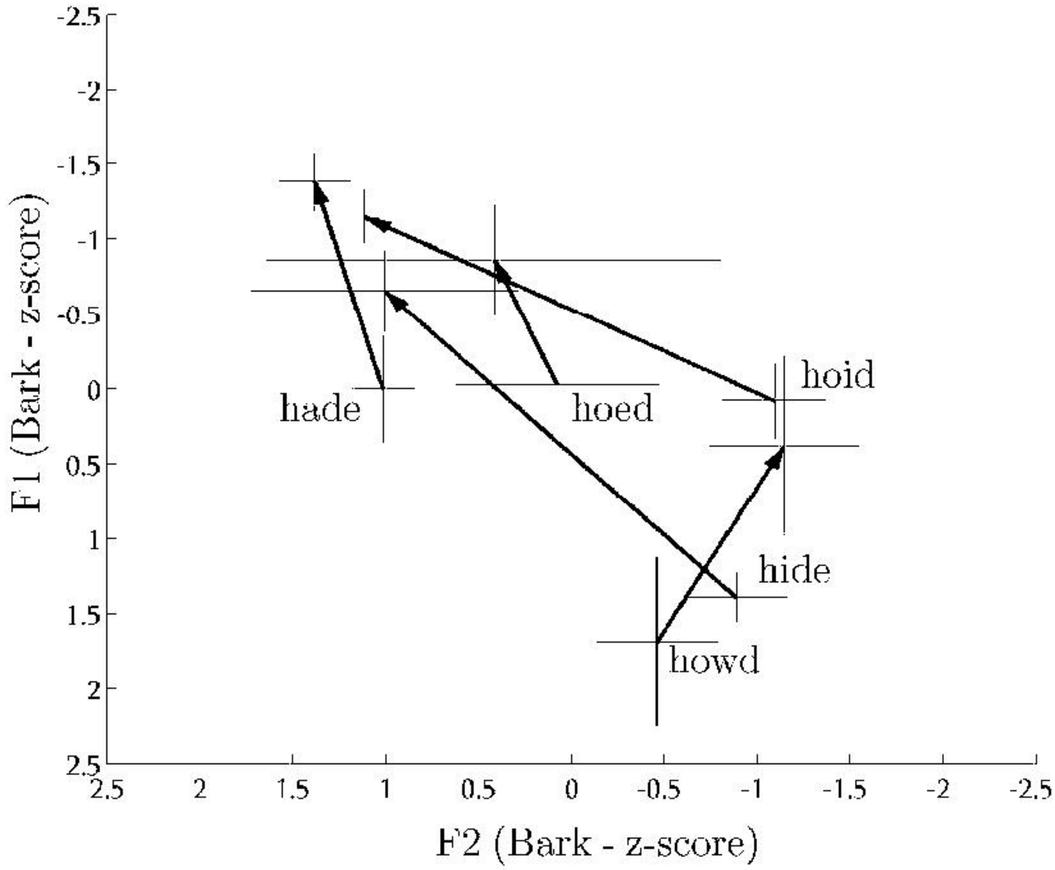

Figure 2-7: Median starting and end point for the 5 closing diphthongs produced by six male speakers of SSBE, adapted from Ferragne and Pellegrino (2010).

In general, SA lacks diphthongs in its vocalic system, therefore, the SA participants of the present study are predicted to have more difficulty perceiving and producing English diphthongs than monophthongs. This difficulty could manifest in replacing some diphthongs with their closest L1 long monophthongs, or producing them with a diphthongal trajectory that is different from the natives'. Since Arabic is a rhotic language, centring diphthongs are predicted to be particularly monophthongal and non-native-like.

2.5 L2 Studies (where Arabic is L1 and English is L2)

Pronunciation problems of Arab learners of English are well documented in the literature of second language research (Flege and Port 1981, Joseph and Odisho 2005, Smith 2001). Smith (2001: 196) suggests that having an Arabic accent in English might be identified due to several phonetic cues. First, he argues that Arabic has more stressed syllables, fewer clear vowel articulations, and frequent use of a glottal stop before initial vowels which leads to breaking up the natural concatenations of English. However, it was unclear which Arabic variety he was referring to, because different varieties have different stress patterns (e.g. Egyptian Arabic and Damascene Arabic). Thus, exploring, in detail, the source of those problems and the possible solutions to overcome or reduce them could provide teachers of English to Arab native speakers with the necessary information to guide the Arabic learners to reduce the Arabic accent in their English productions (Asfoor 1982).

Evidently, most of the difficulties of L2 learners have been argued to be due to the transfer of their L1 forms and structures into their L2 productions. Lado comments that:

> Individuals tend to transfer the forms and meanings, of their native language and culture to the foreign language and culture - both productively when attempting to speak the language and to act in the culture and respectively when attempting to grasp and understand the language and the culture practiced by natives (1957: 2).

However, errors in L2 productions cannot be explained by L1 transfer solely, it is far more complex than it has been assumed in the Contrastive Analysis Hypothesis (CAH), see section 2.2 above. Later L2 learning models (Perceptual Assimilation Model (PAM) and Speech Learning Model (SLM)) have shown the systematicity of L1 patterns transfer into L2 patterns. These models have explored various factors which have been shown to have a significant effect on L2 learning such as Age of Learning (AOL), Age of Arrival (AOA), and amount of L1 and L2 use. Nevertheless, since L2 learners, particularly beginners, do not have access to the L2 allophonic and phonotactic structures, the internal phonological structure of L1 and the phonological processes of

L1 are the learners' main source. Therefore, examining their effect on L2 learning is required and needs further research.

The first step to deal with the pronunciation problems of L2 learners is to perform a comparative study between sound qualities in the target language (L2) and the native language (L1). Language teachers could benefit from such a comparative study in teaching and error correction. Providing quantitative criteria, hand in hand with the qualitative comparison, could also be used to assess students' L2 pronunciation skills (Yang 1996).

Exploring L2 productions of Arab speakers of different dialectal background might give different results. Thus, researchers need to examine the different dialects separately (Newman and Verhoeven 2002). Some researchers of L2 acquisition (Munro 1993) drew their conclusions based on grouping speakers of different Arabic dialects together; and this might lead to questioning the generalisability of their results. For example, Munro (1993) used 21 Arabic speakers from seven different dialects.

Generally, L2 researchers focused on examining the areas of difficulty faced by Arabic learners of English. Regarding vowels, most of these difficulties were due to the lack or absence of the intrinsic features of English vowels in the learners' L1, which were reviewed in the previous sections, such as tenseness, voicing effect of following consonants and diphthongisation. L2 learners are predicted to have greater difficulty producing and perceiving these features.

For example, Mitleb (1981) claimed in his study on Jordanian Arabic that Jordanian speakers could not perceive and produce tense vs. lax distinctions of English vowels. Thus, they transferred Arabic short vs. long vowel duration patterns to English tense vs. lax pairs, and as a result showed significantly larger ratios of tense to lax vowel durations than English speakers did. Thus, 'tenseness' feature has been replaced by the 'length' since length has the main role in the phonemic distinction in Arabic. Accordingly, adult L2 learners did not acquire 'tenseness' feature and replaced it with their L1 length distinction.

The duration of English vowels is also affected by the voicing of the following and preceding consonants. For example, the vowel is longer in English 'tab' than in 'tap'. This feature was also investigated in L2 productions of Arabic learners of English. Mitleb (1982) claimed that voicing effect on vowel duration is not universal since the productions of English minimal pairs spoken by eight Jordanian speakers did not exhibit any temporal implementation due to voicing. A similar test was run on their L1 productions, and the results were as predicted. In Arabic, voicing of the adjacent consonants has no effect on the duration of the vowel. This finding is argued to be due to the fact that L1 speakers focus on producing the long vs. short distinction rather than amending vowel duration depending on the voicing of the adjacent consonants.

Learners of other languages such as Japanese, which has a phonemic length distinction similar to Arabic, have been shown to transfer their L1 experience into their L2 productions. Ingram and Park (1997) examined Japanese and Korean L2 production and perception of Australian English monophthongal non-back vowels /iː ɪ e æ aː/. First, one experienced and one inexperienced group of L2 learners and a group of native English speakers were asked to identify English vowels using a forced-identification task. The same participants were asked to produce a word list of /hVd/ words that included the same vowels they previously identified. Then listeners were asked to transcribe the English stimuli in their native language orthography. They were also asked to rate how closely each stimulus vowel matched the native vowel or representation they chose in their transcription.

The results showed that Japanese subjects transferred their L1 phonemic long vs. short vocalic contrast into their English tense vs. lax vowel productions. On the other hand, the Korean subjects' productions of English tense vs. lax vowels were phonetically graded since quantity does not have a major phonemic role in their native language. Their findings support the SLM hypothesis that L2 learners employ the mechanisms used for establishing the phonetic categories for L1 sounds to acquire L2 sounds. The L2 learners extended their L1 mechanisms which they use to produce or perceive length distinction into their L2 productions.

The Syrian Arabic (SA) participants of the present study are predicted to exhibit some, if not all, of the above quantitative and qualitative difficulties in their L2 perception and production of English vowels. In order to relate the participants' areas of difficulties to their L1 system, a full analysis of the SA vowel system is required, which is the aim of Chapter 4.

Based on the intimate perception-production link suggested by the above gestural theories, physiological and perceptual accounts of speech are needed to understand how speech is generated and interpreted {Scobbie, 2007 #165. The following section presents, in particular, a review of source-filter theory, which is widely used to account for vowels production, which is thoroughly examined in the present study. The following section also presents some methodological issues related to the acoustic analysis of vowels.

2.6 Methodological issues: Acoustic Analysis of Vowels

The previous sections presented a review of several theories of L1 and L2 speech perception with focus on gestural theories. In contrast, fewer attempts were made to develop theories to explain speech production, which might be due to the physiological nature of speech production. The only and widely used source-filter theory, which was first proposed by Fant (1960), was developed to describe vowel production. Since the main interest of this thesis is vowels perception and production, a detailed review of this model is presented in this section.

This model of speech production assumes a **source** for sound production and a **filter** which shapes that sound. The vibrations of the vocal folds at the larynx represent the source of a wave which passes through a filter consisting of the resonating cavities of the vocal tract, between the larynx and the lips. In addition, the model assumes that the source and the filter are independent. This independence means that the actual pitch (i.e. the vibrations of the vocal folds at the larynx) of the sound can be measured separately from the filter resonances. Figure 2-8 below shows a view of the source-filter theory of a vowel produced at 100Hz fundamental frequency (adapted from Ladefoged (1996: 104)).

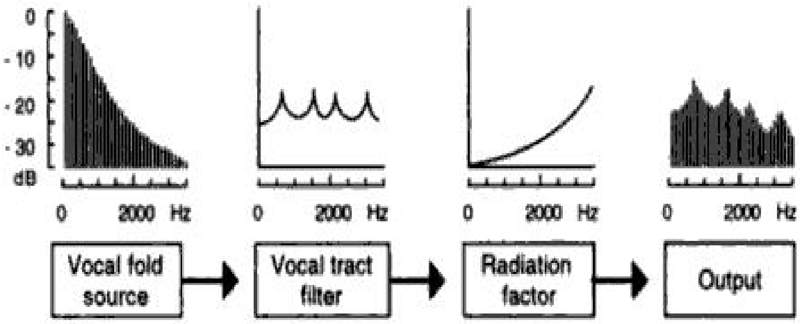

Figure 2-8: Source-filter view of a vowel, adapted from Ladefoged (1996: 104)

Different configurations of the vocal organs result in different pitches or formants. The frequency of these formants has been found to depend mainly on three factors: the backward and forward movements of the tongue, the upward and downward movements of the tongue, and lip position (Ladefoged 1996: 103, 108). The characteristics of these overtone pitches, which are known as formants, determine vowel quality. Accordingly, measuring the frequencies of formants is taken as the typical method used by phoneticians to describe vowels (Ladefoged 2006: 21).

The first two formants, in particular, have been found to distinguish one vowel from another. Stevens (1989) shows that manipulating the first two formants of vowels yields changes to vowel quality. The first formant, F1, is inversely associated with vowel height. That is, F1 decreases as vowel height increases. F2, on the other hand, is inversely associated with place of articulation. F2 decreases as the speaker moves from front to back vowels. Also, the third formant plays a role in determining vowel quality; however, its role is not as clear as the role of the first two formants (Ladefoged 2006: 182, 188). For most purposes, formant frequencies give sufficient information about the acoustic correlates of vowel quality and help in identifying its phonemic category within a particular language.

The frequencies of F1 and F2 in the central most steady state portion of the vowel are traditionally viewed as the principal determinants of vowel quality. However, in coarticulated vowels, formants often do not reach the steady state values of their isolated counterparts, and frequently do not contain any region which can be described as 'steady-state' (Andruski and Nearey 1992). Thus, the best way to extract the frequency values that reflect vowel identity is to examine vowels in isolation or in a relatively neutral context such as /hVd/ (Stevens and House 1963). Paolo et al. (2011: 88) point out that many studies use /hVd/ as a neutral context for vowel productions as the active articulators will be at rest during the production of /h/, which will be produced without any specific ‘tongue body shape, lip protrusion, or constriction in the supralaryngeal cavity (2011: 88)’.

Additionally, vowel duration has been found to affect both vowel quality and quantity (Rosner and Pickering 1994). The effect of vowel duration depends on the speaker’s language background. For example, vowel duration is well known to have a phonological role in vowel identification for languages such as Arabic and Spanish whereas in British English it has an allophonic role that is independent from vowel identity (Kondaurova and Francis 2008, Mitleb 1984b: 23).

In order to have a reliable analysis of vowel quality and quantity, an instrumental phonetic analysis is required since listeners’ perception is not capable of providing this accurate level of analysis (Paolo, et al. 2011: 87). Nowadays, there are many computer programs that can be used to analyse speech sounds into their components. A speech spectrogram is the classic display of analysed sounds. It shows time on the horizontal axis and frequency on the vertical axis, whereas the acoustic energy (intensity) is shown by the degree of darkness (Denes and Pinson 1993: 147, Ladefoged 2006: 185). Figure 2-9 below shows the spectrogram of the phrase ‘say hard again’ as produced by a female English speaker.

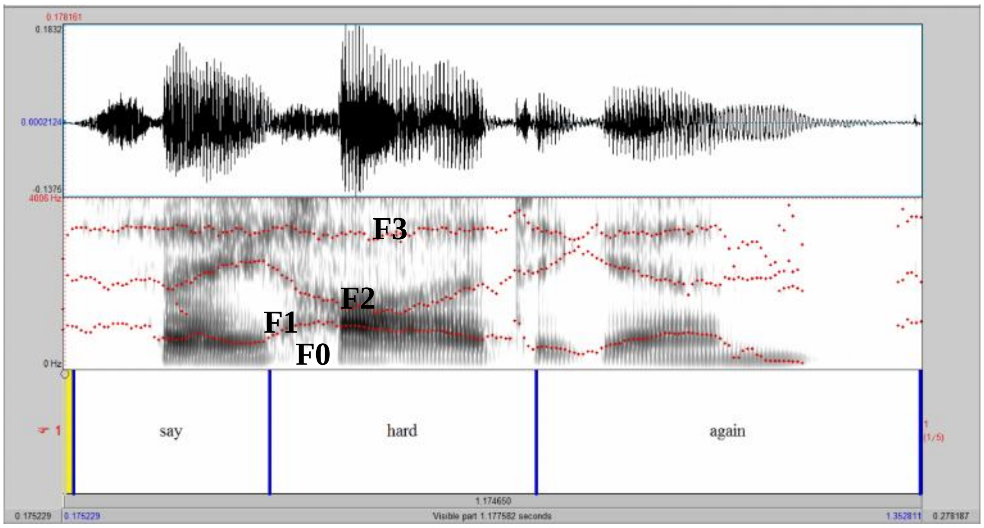

Figure 2-9: Spectrogram showing frequency on the vertical axis, time on the horizontal axis, and intensity by the degree of darkness, of the phrase 'say hard again' as produced by a female English speaker.

Generally, the most reliable method of describing the spectrum of vowels is to set the frequencies up to 4KHz, which is the main frequency range of hearing (Denes and Pinson 1993: 96). Stevens (1989) specifies two reasons for this limit. First, listeners may be insensitive to acoustical details above 4KHz as the auditory resolution gets poor. Second, the acoustic signal becomes weaker at these higher frequencies, and the changes in frequencies at these levels do not have a significant effect on the spectrum of vowels as do the changes in lower formants.

From a physical perspective, the vocal tract of women is generally shorter than men's; and the vocal tract of children is shorter than women's. These differences alter the frequencies of the same phonological vowels produced by male and female speakers of the same language. Generally, females' shorter vocal tract would manifest acoustically in the F2/F1 plane as more dispersed and farther from the males' F2/F1 plane (Rosner and Pickering 1994: 15). Fant (1960) suggested that the formants of women are, on average, 17% higher than the formants of men, and the same formants are even higher for children. Nevertheless, differences in formant frequencies are smaller compared to differences in fundamental frequency. The fundamental frequency (F0) at which the vocal folds vibrate also differs; on average it is approximately 120Hz for men, 220Hz for women, and 330Hz for children (Fant 1960: 87). These frequencies are based on languages such as standard RP and German; other languages might show different patterns. Also, individual differences in the habitual settings of their organs, and language and dialectal differences affect the frequencies of vowels.

Therefore, it is important to take these physiological differences into consideration in any study which involves taking measurements of different gender or age groups. A number of normalisation techniques and methods have been suggested to minimise these biological differences and retain the genuine differences in the speech signal.

2.6.1 Normalisation

The speech signal varies due to several factors. Ladefoged and Broadbent (1957) mentioned three types of variation: phonemic, sociolinguistic and anatomical. As a result of inevitable effects on vowel frequencies, a number of mathematical algorithms were proposed to normalise formant data.

Normalisation can be defined as a transformation procedure used by researchers to compare acoustic data (formants) of different speakers (Adank et al. 2004, Fabricius et al. 2009, Halberstama and Raphael 2004, Pisoni 1997). Such transformation or normalisation methods can be a valuable tool to facilitate across-speaker and across-language comparisons (Yang 1996).

Pisoni (1997) suggests that listeners retain fine phonetic details in the acoustic signal and use other nonlinguistic sources such as: talker's voice, gender, dialect and others, in their speech perception. Thus, he concludes that it is important not to eliminate or reduce this source of variability. Fabricius et al. (2009) state that any normalisation procedure of vowels should aim:

- to eliminate variation caused by physiological differences among speakers;
- to preserve sociolinguistic/dialectal/cross-linguistic differences in vowel quality;
- to preserve phonological distinctions among vowels;
- to model the cognitive processes that allow human listeners to normalise vowels uttered by different speakers.

Choosing the right normalisation method depends on the purpose of normalisation i.e. that is normalising for phonemic, sociolinguistic or anatomical variation. According to Adank et al. (2004) and Fabricius et al. (2009), who compared the efficiency of several normalisation procedures to reduce anatomical variation but reserve phonemic and dialectal information, the LOBANOV procedure succeeds in preserving interesting phonemic and sociolinguistic variation and minimising anatomical variation. LOBANOV is a speaker intrinsic and vowel extrinsic procedure. It is based on the calculation of the Z-score transformation for the formant frequencies using the formula in (3). This formula, among others, is implemented in an online software package NORM (Thomas and Kendall 2007).

(3) $F_{n[V]}^{N} = (F_{n[V]} - MEAN_n)/S_n$

$(F_{n[V]}^{N})$ is the Normalised $(F_{n[V]})$ for formant *n* of vowel *V*

$(MEAN_n)$ is the mean value for formant *n* for a speaker

(S_n) is the standard deviation for the speaker's formant *n*

In sum, the purpose of this section was to present a review of vowel production source-filter theory. This section also reviewed how researchers analyse vowel production data. The domain of this theory is production of first language (L1) vowels. The question now is what happens during the production of the vowels of a second (L2) or a foreign language? Especially, learners of a second language use their L1 phonological gestures as input. Thus, based on the DRT, will the acoustic output of L2 productions match that of the learners' L1?

The review of these different methodologies, which were used to acoustically analyse production data, provided a theoretical background for the acoustic analysis and the normalisation procedures used in the present study.

2.7 Summary

In this chapter an overview of the literature of speech perception and production, with a particular focus on vowels, has been presented in Section 2.1. It was important to understand the mechanisms and theories offered for the L1 acquisition before reviewing those used in L2 learning. Speech perception has been discussed with a focus on articulatory theories i.e. the Motor Theory (MT) and the Direct Realist Theory (DRT), which entail that the articulatory gestures are the objects of speech perception. These theories were reviewed and compared in detail as they are theoretically related to the L2 learning model used to analyse the results of the present thesis, i.e. the Perceptual Assimilation Model (PAM), which is based on an articulatory perspective of speech perception.

Section 2.2 presented a review of second language learning models which attempt to account for learners' perception and production patterns of L2 sounds. The Native language Magnet Theory (NLM) accounts for L1 perceptual data from an acoustic perspective with some implications for L2 perception. The NLM theory is relevant to the analysis of the results of the present thesis because the present study examines vowels acoustically. On the other hand, PAM provided testable predictions on the perceptual patterns of L2 learners based on the degree of similarities and differences between L1 and L2 sounds, which was mainly used to analyse the results of the perceptual experiments and to predict the production patterns of the production experiment. Also, the Speech Learning Model (SLM) was particularly relevant as it provided predictions on L2 production, and thus, production-perception link. These predictions were used to analyse the production data and the L2 production-perception link. Finally, the similarities and differences in the claims and predictions of these models have been compared in section 2.2.4.

After that, the vowel system of the first (L1) and second (L2) language of the participants under investigation were reviewed. Section 2.3 presented a description of the Arabic vowel inventory including its main qualitative and quantitative characteristics with the focus on the SA vowel system, which is the Arabic dialect investigated in this thesis. Then, in section 2.4, a description of the vowel inventory of the target L2 language of the present study, i.e. SSBE, was presented. Similarly, the qualitative and quantitative characteristics of SSBE vowel system were reviewed.

Section 2.5 reviewed L2 studies which examined the production and perception of English vowels by Arabic learners. This section focused on the areas of difficulties which were examined in the literature, and those which are predicted to be found in the present study. The review revealed that little research has been conducted on Arabic vowels, in general, and on SA, in particular. Additionally, very few studies have dealt with L2 production and perception of Arabic learners of English. Thus, the present study aims to fill the gap in the literature by examining the vowel inventory of SA (in Chapter 4) and by examining L2 perception (in Chapter 5 and 6) and production (in Chapter 7) of SA learners of English. In order to begin this exploration, Chapter 3 presents a pilot study which will examine the perception and production of a full list of SSBE vowels by SA participants compared to a full list of their L1 SA vowel system.

Finally, In terms of vowel production, section 2.6 the Source-Filter Theory, which is the only and widely used theory, has been reviewed in detail, followed by the acoustic analysis of vowels including the normalisation procedures used to normalise frequency data. This section particularly helped in determining the appropriate methods to use to collect vowel and L2 perception and production data for this study and the factors that should be taken into consideration such as normalising for gender and language.

3 The pilot study

The present chapter reports the methods and results of a pilot study. The pilot study was conducted since no previous research has been devoted to examine Syrian Arabic (SA) vowels as a first language (L1) or the difficulties SA speakers have when learning a second/foreign language such as English. The pilot study aimed to narrow down the focus of the main study by identifying the English vowels which were difficult to perceive and/or produce by SA learners of English. The Perceptual Assimilation Model (PAM) (Best 1994, 1995, Best and Tyler 2007) framework was used to analyse the data and to set English vowel categories which were difficult or problematic to SA learners. After that, some of the English vowel categories of the pilot study were used as the target of further examination and analyses in the main study.

The perception of English vowels by SA learners was examined in two phases. First, SA participants were asked to categorise English vowels into their closest equivalents among SA vowels in a Perceptual Assimilation Task (PAT) (Gilichinskaya and Strange 2010, Strange et al. 1998). Second, SA participants were asked to identify English vowels in terms of English vowel categories, compared to Native English (NE) participants. After that, English vowel productions of the SA participants were examined and compared to both their L1 (SA) vowel productions and NE productions.

In order to achieve the goal of narrowing down the focus of the main study, the full phonemic vowel inventory of both languages: Standard Southern British English (SSBE) and Syrian Arabic (SA) were included in the pilot study. The vowels included SSBE monophthongs (/iː/, /ɪ/, /ɛ/, /æ/, /ɜː/, /ʌ/, /ɑː/, /ɒ/, /ɔː/, /ʊ/, /uː/) and diphthongs (/eɪ/, /əʊ/, /aɪ/, /aʊ/, /ɔɪ/, /ɪə/, /ɛə/) (Wells 1986), and SA vowels (/iː/, /eː/, /aː/, /oː/, /uː/, /i/, /e/, /a/, /o/, /u/ and /ə/) (Cowell 1964).

In addition to identifying the English vowels which are problematic to SA learners, the findings of the pilot study shaped the main study of the present thesis. For example, the pilot study suggested that rather than being sceptical about the status of SA vowels (particularly mid and central ones), a full acoustic analysis of the SA vowels would help predicting and explaining the participants’ L2 patterns. It was found that including the Arabic pharyngealised allophone [aːˤ], which contrasts with Arabic plain vowel /aː/, was needed in the main study to explain the perceptual and production patterns of a similar English vowel contrast (TRAP, BATH). Another benefit of conducting a pilot was identifying the methodological problems and difficulties. As a result, the identification task of the main study was redesigned to focus on the discrimination and paired-identification of particular English vowel contrasts. Finally, the pilot study could not have been satisfactory on its own since it was conducted with few participants, who were not an exact match to the target learners of the present thesis.

The rest of this chapter will present the methodology of the three experiments described above: The Perceptual Assimilation Task (PAT), The Identification Task, and The Production Task. The presentation of each experiment includes hypotheses, stimuli, recording procedures, participants, stimulus material, and data analysis procedures. After that, this chapter is concluding with a summary of the main findings, limitations and concluded remarks for the main study.

3.1 The Perceptual Assimilation Task

The PAT was designed to examine the perception of English vowels by SA learners of English. In this task, the perception of a full list of English vowels was examined in terms of a full list of SA/L1 vowels. Generally, the PAT aimed to explore the following research questions:

RQ1: How do SA participants perceptually assimilate English vowels to the five vowel qualities of their SA/L1 phonological inventory?

- How do SA participants perceptually assimilate the English tense vs. lax vowels to the long vs. short vowels of Syrian Arabic?
- How do SA participants perceptually assimilate English monophthongs vs. diphthongs to the vowels of SA?

RQ2: To what extent do PAM perceptual assimilation categories account for the participants' perceptual assimilation of English vowels to their L1 phonemic vowel inventory?

3.1.1 Predicted Assimilation Patterns

PAM determines the assimilation patterns of L2 learners based on articulatory-phonetic similarities between the L1 and the target language L2 (Best, et al. 2001). Following Best and her colleagues, the similarities and differences between SSBE vowels and SA vowels were determined based on their articulatory-phonetic characteristics. Table 3-1 below presents SSBE vowels and their predicted SA equivalent vowels; it also presents the possible articulatory-phonetic similarities between these groups of sounds.

Table 3-1: SSBE vowels and their predicted equivalents in SA based on their articulatory-phonetic similarities

SSBE vowels	SA equivalents	Articulatory-phonetic similarity
(FLEECE, NEAR)	/iː/	high, front, long vowels
(DRESS, NURSE, SQUARE, FACE)	/eː/	mid, front-central area, long
(TRAP, BATH)	/aː/	low, long
(THOUGHT, GOAT)	/oː/	mid, back, long
GOOSE	/uː/	high, back, long
KIT	/i/	mid-high, front, short
DRESS	/e/	mid, front, short
STRUT	/a/	mid-low, central, short
LOT	/o/	back, short
FOOT	/u/	mid-high, back, short
PRICE	/aj/	diphthong, vowel-gilde sequence, similar onglide and offglide components
MOUTH	/aw/	diphthong, vowel-gilde sequence, similar onglide and offglide components
CHOICE	/oj/	not similar to any SA vowel, but might be perceived as vowel-glide sequence

As shown in the table, most English long vowels and diphthongs are predicted to be assimilated into SA long vowels. However, English PRICE and MOUTH are predicted to map to SA vowel-glide sequences /aj/ and /aw/ which have similar onglide and offglide components. For this reason, the two Arabic words /ʔaj/ ‘which?’ and /ʔaw/ ‘or’ were included in the PAT. The English CHOICE diphthong is a special case distinct from the previous diphthongs. It is not similar to any SA vowel or vowel-glide sequence. Nevertheless, it might be treated like the English PRICE and MOUTH diphthongs and be perceived as a vowel-glide sequence of /oj/, which has not been reported as a possible sequence in SA, rather than a single vowel. English short vowels, on the other hand, are predicted to map to SA short vowels.

3.1.2 Recordings

Two SSBE speakers were recorded in an acoustically treated recording area in the linguistic lab in the Department of Language and Linguistic Science at the University of York. The recordings were carried out using Behringer C1 large diaphragm condenser microphone, TAC Scorpion-Mixing desk, M-Audio LT1010 PCI Audio card, and Adobe audition v1.0 on PC software with 16bit and 44.1 KHz sampling rate.

One SA 36 year old male speaker produced SA alternative target options to be used as responses in the PAT. He was recorded in a quiet room using Marantz Professional Portable Solid State Recorder, PMD660 model. The microphone was Shure Professional unidirectional, which was a head-worn dynamic microphone. Audio files were recorded at 44.1 KHz sampling rate as .wav files on a compact flash TM memory card.

3.1.3 Participants

One male and one female SA participant participated in the pilot study PAT. Due to dialectal variation in Arabic, only Syrian Arabic speakers of the Damascene variety were chosen to participate. The SA participants were postgraduate students in the UK, and they received English formal education during school and university between 12 and 18 years in Syria. A questionnaire was used to obtain information about the participants, in general, and their language background, in particular (see Appendix A).

Table 3-2: Summary of the language background questionnaire obtained from the Syrian Arabic participants and the scores of their English skills

Participant	Age	Years of Formal education	Years in UK	Speaking	Understanding	Reading	Writing	Average score
SA1	29	12	3	6	6	7	7	6.5
SA2	26	18	1	6	7	7	7	6.7
Average	27.5	15	2	6	6.5	7	7	6.6

Table 3-2 above shows that the average age of Syrian Arabic participants is 27.5 years old. The average of their English formal education is 15 years, and on average they spent two years in the UK. The participants were asked to rate their own language skills abilities on a 7 points scale (1- very poor to 7-excellent). The ratings of their English skills are high in general, however, the rate for reading and writing was higher than the rates for speaking and understanding; this was an expected result since their English input relied mostly on formal education which focused mostly on reading and writing.

3.1.4 Stimulus Material

The English stimulus material for this study consisted of a set of monosyllabic /CVC/ words. For each of the English vowels (/iː/, /ɪ/, /ɛ/, /æ/, /ɜː/, /ʌ/, /ɑː/, /ɒ/, /ɔː/, /ʊ/, /uː/, /eɪ/, /əʊ/, /aɪ/, /aʊ/, /ɔɪ/, /ɪə/, /ɛə/), 4 or 5 representative real English words were chosen, see Table 3-3 for a list of the words for each vowel. The target words were put in the phrase ‘say ___ again’. SSBE speakers were asked to read three randomised blocks of the target words. Thus, the speakers produced 85 items × 3 repetitions = 255 items.

Table 3-3: List of English monosyllabic words representing English vowels which were used as stimuli for the PAT

SSBE vowel	Word	SSBE vowel	Word
FLEECE /iː/	Heat, beat, sheet, keep, deep	FACE /eɪ/	Faith, fate, case, face, take
KIT /ɪ/	Hit, bit, kit, pick	GOAT /əʊ/	Coat, hope, soap, goat, boat
DRESS /ɛ/	Bet, debt, net, set, met	PRICE /aɪ/	Sight, bite, fight, bike, kite
TRAP /æ/	Hat, cat, fat, sat, bat	MOUTH /aʊ/	House, mouth, doubt, mouse, shout
STRUT /ʌ/	Cut, shut, but, cup	CHOICE /ɔɪ/	Join, coin, boys, toys, voice
BATH /ɑː/	Tart, dark, heart, part, park	NEAR /ɪə/	Tear, near, hear, dear, fear
LOT /ɒ/	Hot, shot, pot, dot	SQUARE /ɛə/	Hair, dare, care, bear, their
THOUGHT /ɔː/	Bought, caught, short, taught		
FOOT /ʊ/	Put, foot, cook, book, took		
GOOSE /uː/	Boot, tooth, suit, soup, shoot		
NURSE /ɜː/	Purse, shirt, hurt, nurse, curse		

Then, two lists consisting of the best production for each vowel were prepared for two SSBE speakers separately. The best productions were chosen by the researcher to have no hesitation or mispronunciation. The two lists (120 items) were judged as being good instances of their intended category, by another two native SSBE speakers (Tsukada et al. 2005). The first judge had some knowledge of phonetics and the second was a phonetician. The judges were asked to give a rate on a five point scale for each word (1-it does not represent its vowel category to 5-it represents the intended vowel category). Finally, the productions of the speaker, with the highest points, a 41 year old male English speaker from Oxford, were randomised and presented to SA participants.

Following Strange (1998), the SA alternative responses consisted of a full list of L1 vowel categories (/iː/, /eː/, /aː/, /oː/, /uː/, /i/, /e/, /a/, /o/, /u/ and /ə/). Thus, the SA alternative responses consisted of 11 monosyllabic /CVC/ words and disyllabic /CVCC**V**C/ words for short vowels since they do not occur in monosyllabic words. SA sequences of vowels and glides /aj/ and /aw/, which are diphthong-like sounds, were also added to the list because some English diphthongs were predicted to map to them. Together the SA alternative responses consisted of 13 alternatives; see Table 3-4 for a list of the Syrian Arabic alternative words. Since these responses were the same throughout the whole task, the large number of responses (13) did not cause any difficulties to participants.

Table 3-4: A list of the SA alternative responses used in the PAT

SA Vowel	Arabic Word	Transcription	Gloss
1. /iː/	تين	/tiːn/	figs
2. /ɪ/	دب	/dɪbb/	bear
3. /eː/	سيف	/seːf/ -	sword
4. /e/	مَلِك	/malek/	king
5. /aː/	باب	/baːb/	door
6. /a/	أَب	/ʔabb/	father
7. /oː/	موز	/moːz/	bananas
8. /o/	مشمُش	/mɪʃmoʃ/	apricot
9. /uː/	فول	/fuːl/	beans
10. /u/	كُرة	/kura/	ball
11. /ə/	سِن	/sənn/	tooth
12. /aj/	فَي	/faj/	shadow
13. /aw/	ضَو	/dˤaw/	light

SA alternative words were selected with the help of two SA native speakers, who were asked to judge whether the SA words selected by the researcher represent the SA vowel categories suggested by Cowell (1964). The final set of SA words was presented to SA participants in conventional Arabic orthography simultaneously with the presentation of the audio SSBE recordings.

In Arabic, short vowels are used as diacritics, which are only used with carefully written formal speech. Arabic speakers usually do not need short vowels as they can infer what they are from the context. As a result, without an appropriate context the participants of the present thesis might not be able to produce the required target short vowels and mix them. To overcome this problem, pictorial alternatives were chosen to present the picture of the L1 alternative in addition to its spelling. Also, audio productions of the same L1 alternatives, which were produced by the SA male speaker, were presented simultaneously with the visual and written words; and the participants were allowed to play the target SA alternatives to listen to if they wished during the task. The listeners were familiarised with the alternative responses before undertaking the perceptual test, and the target syllables were highlighted.

3.1.5 Procedures: the listening experiment

In the PAT, the SA participants were informed that they were going to listen to SSBE words mentioned in section 3.1.4 above, and that these words will be presented in the phrase 'say ___ again'. The experiment was run using a Praat MFC experiment (Boersma and Weenink 2009). The participants were directed to listen to the target English word and then select the SA/L1 vowel category to which each English vowel was most similar. There were 13 L1 alternative audio responses, which listeners could listen to as many times as they wished by clicking on the relevant button, as in Figure 3-2.

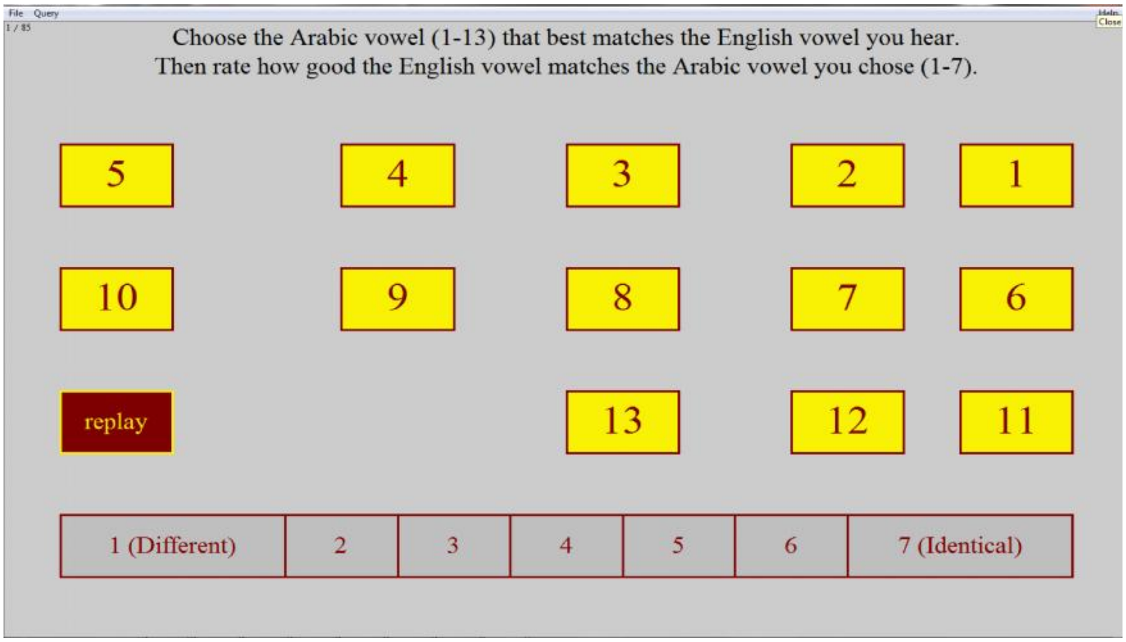

Figure 3-2: a screen shot of the PAT showing the 13 audio buttons of the SA alternative responses given in Table 3-4 above

Figure 3-2: a screen shot of the PAT showing the pictures and written form of the 13 SA alternative responses given in Table 3-4 above

The visual representations of the SA alternative words with the written forms were presented on a separate sheet (as in Figure 3-2 above), so participants could see them simultaneously while listening to their production. The participants were hearing and seeing whole SA words in a citation form and not just the vowels. Before moving to the next item, participants were asked to rate its category goodness of fit as a good example of the SA/L1 vowel category on a 7-point scale. This scale reflected the perceived similarity between the English and SA/L1 sound: 1-3 were treated as different, 4- different but shared some properties with L1 vowel category, and 5-7 similar. The participants were given a replay button to listen to the target English word as many times as they wished. Once they decided and gave the category goodness of fit rating for their choice, they pressed the OK button and moved to the next item.

3.1.6 Data Analysis and Results

The classifications/categorisations of each participant were pooled into a confusion matrix of all English and SA vowels (i.e. a table that includes all possible categorisations between SA and SSBE vowels). After that, these categorisations were analysed using PAM (Best 1995) assimilation categories. Additionally, these categorisations were analysed according to their temporal and spectral properties (Strange, et al. 1998).

The number of responses of the SA participants was calculated for each English vowel. Table 3-5 below presents the confusion matrix of the vowels under investigation; it presents the percentages and average goodness of fit ratings for the categorisation of each English vowel. It also shows the total number of responses for each vowel; as can be seen the numbers are too small to run statistical analyses. Thus, only descriptive presentation of the results will be presented.

Table 3-5 shows that most English vowels are categorised quite clearly to particular SA phonological vowel categories, such as English (FLEECE, DRESS, TRAP, STRUT, BATH, THOUGHT, FOOT, GOOSE, FACE, PRICE, MOUTH, SQUARE), and these vowels are categorised mostly with high goodness of fit ratings, i.e. 5 or more. A few English vowels (KIT, LOT, GOAT, CHOICE, NEAR) do not show clear categorisation patterns as the previous vowels, and their goodness of fit ratings are also lower, i.e. 4 or less.

Table 3-5: percentages of the categorisation of English vowels into SA vowels, showing goodness of fit ratings (in brackets: 1 different - 7 identical), and the total number of responses for each English vowel (in bold). Shaded cells show the highest categorisation tendency. English vowels in bold are the ones included in the main study in the following chapters.

	/iː/	/i/	/eː/	/e/	/aː/	/a/	/oː/	/o/	/uː/	/u/	/ə/	/aj/	/aw/	**NUM**
FLEECE **/iː/**	100 (6.1)													**10**
KIT **/ɪ/**	13 (3)	50 (5)		37 (4)										**8**
DRESS **/ɛ/**			20 (6)	70 (5.5)							10 (6)			**10**
TRAP **/æ/**					60 (5.3)	40 (4.5)								**10**
STRUT **/ʌ/**						88 (5.5)		12 (6)						**8**
BATH **/ɑː/**					100 (5.5)									**10**
LOT **/ɒ/**						13 (7)	25 (5)	50 (5.5)		12 (5)				**8**
THOUGHT **/ɔː/**							88 (5.4)	12 (5)						**8**
FOOT /ʊ/									10 (7)	90 (6.4)				**10**
GOOSE /uː/									90 (6.2)	10 (5)				**10**
NURSE /ɜː/			40 (4.7)	10 (4)			20 (3)				30 (4.6)			**10**
FACE **/eɪ/**			60 (6.1)	20 (6)								20 (6)		**10**
GOAT **/əʊ/**						12 (1)	38 (5)	25 (6)		25 (5.5)				**8**
PRICE /aɪ/												100 (4.1)		**10**
MOUTH /aʊ/													100 (5.9)	**10**
CHOICE /ɔɪ/							40 (3.5)		10 (1)				50 (1)	**10**
NEAR **/ɪə/**	50 (4)	30 (3.6)	10 (3)	10 (3)										**10**
SQUARE **/ɛə/**			70 (5.7)	20 (4)							10 (6)			**10**

A closer look at the categorised English (/æ/, /ɑː/), (TRAP, BATH henceforth), vowels shows that both are categorised into a single SA/L1 long phonological category /aː/ and both are rated as good exemplars of the L1 category. Thus, this L2 contrast belongs to PAM's SC (Single Category) assimilation type. Similarly, English (/eɪ/, /ɛə/), (FACE, SQUARE henceforth), vowels are categorised into a single SA/L1 long phonological category /eː/ and they are equally rated as good exemplars of the L1 category. Thus, this L2 contrast also belongs to PAM's SC assimilation type.

On the other hand, English /ɪ/, (KIT henceforth), is categorised mostly to SA /i/ with a relatively high goodness of fit rate. However, a great number of English KIT are categorised as SA /e/ with goodness of fit rating (4), which shows that English KIT shares some properties with SA /e/ though it is categorised mostly as SA /i/. In this case, English (/ɪ/, /ɛ/), (KIT, DRESS henceforth), contrast matches the description of PAM CG (Category Goodness) assimilation type because both are categorised as SA /e/ with English DRESS being closer to SA /e/ in terms of the number of responses 70% and the goodness of fit rating (5.5).

English /ɒ/, (LOT henceforth) is categorised into SA back vowels (particularly /o/) with relatively high goodness of fit ratings. However, it is categorised 13% into SA front low vowel /a/ with identical (7) goodness of fit rate. As can be seen in the table, English /ʌ/, (STRUT henceforth), is also mostly categorised as SA /a/ with high goodness of fit rate. Thus, the English (LOT, STRUT) contrast also maps to the PAM CG assimilation type because these two vowels are categorised as SA /a/ with English STRUT being closer to SA /a/ in the number of responses but not in the goodness of fit rating.

Among English diphthongs, the /ɔɪ/ (CHOICE) diphthong does not have a clear categorisation into any of the SA vowels and it is rated as different from any SA vowel i.e. it was given a goodness of fit rating less than 3. Thus, this vowel can be mapped into PAM's Non-Assimilable (NA) perceptual assimilation type. As a native speaker of SA, my prediction would be that SA participants perceived the English CHOICE diphthong as a sequence of a mid-back vowel and a glide /j/, which is similar to the manner SA participants followed in their categorisation of the English PRICE and MOUTH into SA vowel-glide sequences /aj/ and /aw/, respectively, which can be seen in Table 3-5. However, further examination, with different experimental settings from those used in the present thesis, is needed to confirm this prediction. Therefore, the English CHOICE diphthong will be excluded from the analysis in the main study.

As for the English /ɪə/ diphthong, (NEAR henceforth), its categorisation spreads over SA front vowels with low goodness of fit ratings. However, it shows a tendency of categorisation into the SA high front long vowel /iː/ with a (4) goodness of fit rating, which indicates that English NEAR diphthong is different from SA /iː/ but they share some properties, which might be the quality of the first part of the English diphthong. This also means that SA /iː/ and English NEAR are perceived as phonologically similar but phonetically deviant from each other. According to PAM, the English (NEAR, FLEECE henceforth), contrast maps to Uncategorised- Categorised (UC) assimilation type, in which English FLEECE is categorised as L1 /iː/ vowel and English NEAR is the uncategorised vowel.

Also, the /əʊ/ diphthong (GOAT henceforth) is categorised into SA back vowels, particularly the long mid back vowel /oː/, with relatively high goodness of fit ratings. The English /ɔː/ vowel (THOUGHT henceforth) was also categorised, highly and with a high goodness of fit rate, as SA /oː/. Thus, the English (GOAT, THOUGHT) contrast belongs to PAM's CG assimilation type because the members of the L2 (GOAT, THOUGHT) contrast were categorised as SA long vowel /oː/ with THOUGHT being closer to the SA phonological category /oː/.

Table 3-6 below, summarises the results of the PAT and the analysis of the results according to PAM perceptual assimilation types. As can be seen, among English vowels two vowel contrasts (TRAP, BATH) and (FACE, SQUARE) were considered as PAM SC assimilation type. PAM predicts that initially learners will have difficulty discriminating these L2 vowels. However, learners in later stages could perceptually attune to this contrast and establish a new phonetic category to at least one of the L2 sounds before they would be able to establish a new phonological contrast (Best and Tyler 2007).

Table 3-6 also shows that three English contrasts were considered as PAM CG assimilation type: (KIT, DRESS), (LOT, STRUT), and (GOAT, THOUGHT). The PAM prediction for the CG contrast is that learners will discriminate the L2 contrast but not as well as a TC (Two Category) assimilation type. It also predicts that a new phonetic and phonological category will be eventually formed for the deviant L2 vowels

KIT, LOT, and GOAT whereas DRESS, STRUT, and THOUGHT vowels which were perceived as good exemplars of L1 categories will be phonetically and phonologically perceived similar to L1 category and no new category is likely to be formed. Finally, the English (NEAR, FLEECE) contrast was considered as PAM UC assimilation type. PAM predicts that learners will have little difficulty in discriminating these sounds.

Table 3-6: summary of the results of the PAT including SSBE/ L2 contrasts, their categorisations by SA/L1 participants into their closest equivalents in SA/L1, and their PAM assimilation type.

SSBE/L2 contrasts	SA/L1 equivalent	PAM assimilation type
BATH	/aː/	SC
TRAP	/aː/	
FACE	/eː/	SC
SQUARE	/eː/	
KIT	/i/, /e/	CG
DRESS	/e/	
LOT	/o/, /a/	CG
STRUT	/a/	
GOAT	/oː/	CG
THOUGHT	/oː/	
FLEECE	/iː/	UC
NEAR	/iː/	

Generally, the pilot study PAT provided some answers to the research questions in section 3.1 above, which will be further examined in the main study:

RQ1: How do SA participants perceptually assimilate English vowels to the five vowel qualities of their SA/L1 phonological inventory? The PAT showed that the participants were able to fit all English vowels into SA categories, using any detected quantitative, qualitative, or feature similarities.

- How do SA participants perceptually assimilate the English tense vs. lax vowels to the long vs. short vowels of Syrian Arabic?

The answer would be that the participants, as predicted, mapped the English tense vowels such as FLEECE and TRAP into SA long vowels whereas English lax vowels such as KIT and LOT into SA short vowels. Thus, it can be argued that the participants used their L1 quantitative-based vowel system to perceive English vowels.

- How do SA participants perceptually assimilate English monophthongs vs. diphthongs to the vowels of SA?

Clearly, the participants used their L1 vowel-glide sequences (/aj/ and /aw/) to map English diphthongs such as PRICE and MOUTH. However, they mapped FACE and SQUARE diphthongs into SA mid long vowels, which can be argued to be an extension of their L1 vowel-glide interaction (which led to those mid long vowels originally). As for the CHOICE diphthong, it was more difficult to assimilate to any SA vowel, or vowel-glide sequence, which suggests another path to perceive it; i.e. a possible vowel-glide sequence /oj/ that was not reported in earlier studies.

RQ2: To what extent do the PAM perceptual assimilation categories account for the participants' perceptual assimilation of English vowels to their L1 phonemic vowel inventory? The results showed that PAM accounts for the assimilation patterns of most English vowels, which were mapped into PAM assimilation types as presented in Table 3-6 above. Based on these categorisations, PAM predicts that the members of the CG assimilation type contrasts will have better discrimination than both SC and UC assimilation type contrasts (Best, et al. 2001). In order to examine the discrimination of the English contrasts in Table 3-6 and the predictions of PAM regarding assimilation types, an identification experiment was designed, in which SA and NE participants were asked to identify and discriminate English vowels in a forced identification task. The methods and results of the identification task will be presented in the following section.

3.2 The Identification Task

The pilot study identification task was also designed to examine the perception of English vowels by SA learners of English. The identification of a full list of English vowels was examined in terms of English vowel categories. Generally, the identification task aimed to explore the following research questions:

RQ1: How do SA participants perceive English vowels in terms of the English phonemic vowel inventory compared to NE participants?
RQ2: Which English vowel contrasts are difficult to identify or discriminate by SA participants?
RQ3: Is this difficulty related to quality or quantity overlap?
RQ4: To what extent can PAM assimilation categories account for the results of the present study?

3.2.1 Hypotheses

In general, the focus of the identification experiment was to examine the identification of all English vowels by SA participants compared to NE participants, i.e. that is to determine the English vowel category they are listening to. The identification of the English vowel contrasts which were analysed in the previous task will be examined in particular. The results of the PAT were analysed in the PAM framework, which provided empirical predictions to be examined in the identification task as follows:

- Two English vowel contrasts (TRAP, BATH) and (FACE, SQUARE) were found to belong to PAM SC assimilation type. PAM prediction was that the members of the SC contrasts will be poorly discriminated since each contrast was assimilated into a single L1 category /aː/ and /eː/, respectively, and were equally rated as good exemplars of that category.
- Three English vowel contrasts (KIT, DRESS), (LOT, STRUT), and (GOAT, THOUGHT) were analysed as PAM CG assimilation type. The members of the CG contrasts are predicted to have from moderate to good discrimination since each contrast has been categorised into a single L1 category /e/, /a/ and /oː/ with the second member of these contrasts being closer to the L1 vowel category, in terms of the number of categorised tokens and goodness of fit ratings.
- The English (NEAR, FLEECE) contrast was analysed as PAM UC assimilation type, in which English FLEECE was categorised as L1 /iː/ vowel with a high goodness of fit rate whereas English NEAR was

categorised as L1 /iː/ with low goodness of fit rate. Thus, English FLEECE was considered the categorised vowel and English NEAR was the uncategorised vowel, and these vowels are predicted to have very good discrimination by SA participants.

3.2.2 Recordings

The recording procedures and methods were similar to those used to record the English speakers for the PAT (see Section 3.1.2).

3.2.3 Participants

The same two SA participants who took part in the PAT also took part in the identification task; for more information about the SA participants see Section 3.1.3. As a control group, two Native English (NE) participants took part in the identification task. Similar to the SA participants, a questionnaire was used to extract information about the English participants and their language background. They were chosen from students and staff at the University of York. NE participants were chosen from Southeast London and the Home Counties, and they were 30 years old on average.

3.2.4 Stimulus Material

The stimulus material for this experiment consisted of a list of real monosyllabic /CVC/ words, which included some of the words from the previous experiment. The list was chosen to consist of simple and familiar words to the participants, to avoid misidentification for reasons other than vowel category. All target words were put in the same phrase used in the PAT 'say ____ again'.

The same English male speaker form the PAT produced the stimulus material for the identification task. The material was chosen and judged following the same procedures used in the previous experiment (see Section 3.1.4).

The present study was still a pilot study, designed to explore the perception and production of all English vowels by SA learners, in order to narrow down the focus of the main study. Thus, the identification task examined the identification of all English vowels and not only the English vowel contrasts which were found to have categorisation issues in the PAT. However, the focus of the analysis will be on those vowel contrasts presented in Table 3-6.

The design of the alternative English responses of the identification task was based on vowel height, place, or quantity differences. Every (a. b. c.) group of choices consisted of minimal pairs which differed only in the vowel, and this difference was in terms of vowel height, place, or quantity. For diphthongs, additional criteria were used, such as the starting or the ending point of the diphthong, i.e. the alternatives were chosen to differ in the beginning or the ending of the diphthong. There was also a mix of long vowels and diphthongs in the choices because Arabic does not have diphthongs but does have long vowels. Thus, adding this mix might reveal if SA participants identify some English diphthongs as other English long vowels instead, particularly since SA participants categorised many English diphthongs as SA long vowels in the previous experiment (see Section 3.1.6). All together, participants were presented with 120 items to listen to and identify; see Appendix B for a full list of the alternative responses used in the identification task.

3.2.5 Procedures: the listening experiment

The listening experiment was run in the recording room in the department of Language and Linguistic Science at the University of York. The experiment was run using the Praat MFC experiment (Boersma and Weenink 2009). The participants were familiarised with the English words using spellings and pictures, and if needed, the researcher produced the target words as many times as they required.

Using a forced-choice identification of the stimulus, listeners were asked to choose one of three alternatives presented on a response sheet using English orthography with a. b. and c. choices. This was a self-paced task; the participants were directed to listen to the phrase "say ___ again", choose the word from the response sheet that best matches the target word, and then press the a. b. or c. alternative on the computer screen (as can be seen from the screen shot of the experiment in Figure 3-3 below). A replay button was added to allow listeners to listen to the words as many times as they wish, once they decide they press the OK button and move to the next item.

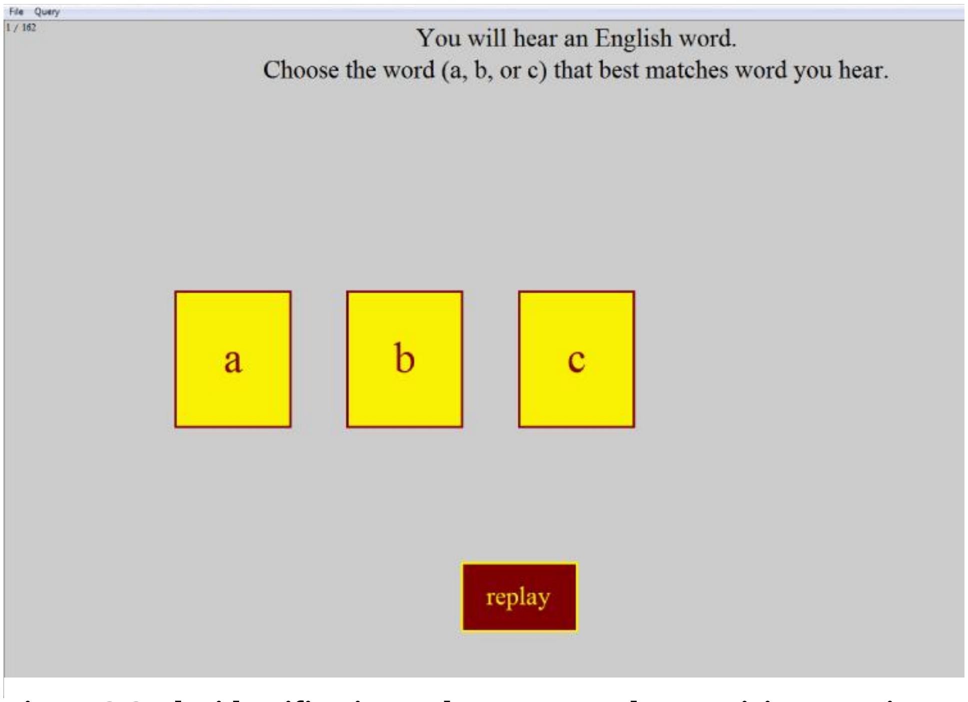

Figure 3-3: the identification task as presented to participants using Praat MFC experiment, this was presented simultaneously with a response sheet with the choices written in English

3.2.6 Data Analysis and results

First, a matrix of errors of identification was made for each listener (a table with all possible identifications for all English vowels). It included a full list of English vowels and all the vowels which were used as distracters, as shown in Table 3-7 below. The shaded cells represent the correct responses i.e. correct identifications and the other cells represent the incorrect responses. This table represents the errors of identification grouped together across the two SA participants. The same table was also made across the two NE participants. Then, the percentages of correct and incorrect identifications were calculated for each vowel and for each group.

As can be seen in Table 3-7, most English vowels were 100% correctly identified by SA participants. Only English (KIT, STRUT, LOT, GOOSE, GOAT, and NEAR) vowels were misidentified by SA participants. Some of these vowels had clear assimilations into SA vowels, but they had misidentifications such as the GOOSE vowel. On the other hand, the assimilation patterns of the other vowels were reflected on their identification. The results of the identifications of NE participants are not presented here because they 100% correctly identified all English vowels.

Table 3-7: matrix table of identification used to extract the errors of identification of the English vowels by SA participants, showing percentages of correct (highlighted cells) and incorrect identifications for each vowel, it shows the number of items presented to participants for each vowel. The results for the six contrasts under investigation are presented first, in terms of PAM assimilation types.

PAM type	SSBE vowel	FLEECE	KIT	DRESS	TRAP	STRUT	BATH	LOT	THOUGHT	FOOT	GOOSE	NURSE	FACE	GOAT	PRICE	MOUTH	CHOICE	NEAR	SQUARE	Num of items
SC	TRAP				100															16
	BATH						100													14
SC	FACE												100							10
	SQUARE																		100	10
CG	KIT		50	50																14
	DRESS			100																16
CG	LOT					6		83	11											18
	STRUT					87						13								16
CG	GOAT											10		90						10
	THOUGHT								100											12
UC	FLEECE	100																		12
	NEAR	9											8					83		12
	CHOICE																100			10
	FOOT									100										14
	GOOSE									7	93									14
	NURSE											100								14
	PRICE														100					10
	MOUTH															100				12

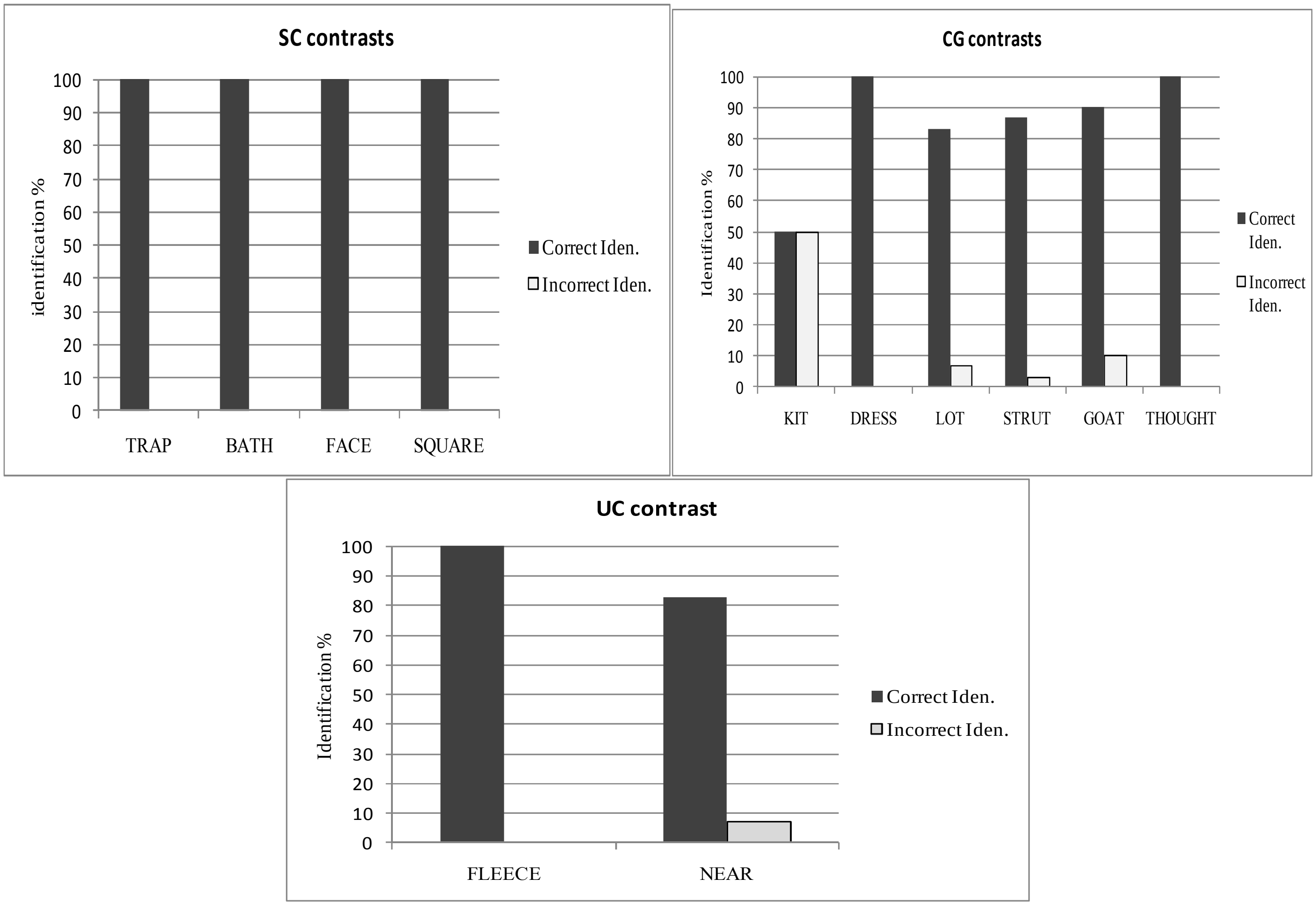

Figure 3-4: percentages of correct and incorrect identification of the English vowels by SA participants, divided according to their PAM assimilation type (SC, CG, and UC)

Figure 3-4 shows the percentages of correct and incorrect identification of the English vowels by SA participants. As can be seen, the members of the two SC contrasts (TRAP, BATH) and (FACE, SQUARE) were 100% correctly identified. According to PAM, however, SC contrasts are predicted to have poor discrimination by SA participants since they were categorised as a single L1 phonological vowel category. The identification and the discrimination of the SC contrasts by learners is argued by Best to depend on the perceived differences between those vowels, but the exact kind of similarities or differences are not specified (Best 1994, Best and Tyler 2007). It seems that SA participants perceive the members of the English SC contrasts (TRAP, BATH) and (FACE, SQUARE) with enough phonetic distance to distinguish them.

According to Flege's (1995: 239) SLM (Speech Learning Model) L1 and L2 sounds exist in a common phonological space, in which L1 and L2 sounds are perceived at the allophonic level. Flege argues that listeners are able to detect phonetic differences between L1 and L2 sounds if the perceived phonetic difference between these sounds is great enough. Similarly, if the perceived phonetic difference between two L2 sounds is great enough, one can argue that learners will be able to detect the phonetic difference between the L2 sounds. In the present case, we suggest that SA participants were able to detect the phonetic difference between English (TRAP, BATH) vowels, because it is very close to the phonetic/allophonic distinction between the Arabic plain vs. pharyngealised (/aː/, [aː^{ʕ}]) vowel contrast.

Similarly, SA participants were able to detect the phonetic difference between the two English diphthongs (FACE, SQUARE), which is mainly in the off-glide of the diphthongs (the FACE off-glide is towards a high front vowel and SQUARE off-glide is towards a mid central vowel). Thus, it can be argued that learners have more options to refer to when they perceive L2 sounds if they use their L1 phonemic categories as well as their L1 phonetic/allophonic categories. In the case of this contrast, the off-glide phonetic component of the diphthongs was used to distinguish two full phonological diphthongal categories. Particularly, SA is argued to have short high vowel /i/ and a central vowel schwa (Cowell 1964) which might be assimilated to the off-glide components of (FACE, SQUARE) diphthongs.

The identification of members of the three English CG contrasts (KIT, DRESS), (LOT, STRUT), and (GOAT, THOUGHT) is more difficult for SA participants. As predicted form the previous experiment, in all cases the English vowel which was categorised as closer to the SA vowel, in terms of the number of responses and goodness of fit rate, was identified correctly more often than the deviant vowel. The most obvious example is the English (KIT, DRESS) vowels which were correctly identified 50% and 100%, respectively. Further analysis of the data shows that the 50% of English KIT misidentification was identified as English DRESS. Thus, the English (KIT, DRESS) contrast was poorly discriminated by SA participants. Similarly, the second members of English (LOT, STRUT) and (GOAT, THOUGHT) contrasts were closer to SA vowels and though the difference was much smaller, they were correctly identified more than their deviant counterparts, see Figure 3-4.

The results of the identification task conform to PAM predictions for the good (LOT, STRUT) and (GOAT, THOUGHT) discrimination but not for the poor discrimination of (KIT, DRESS). A potential explanation for the poor discrimination of English (KIT, DRESS) by SA participants is that the SA vowel system might not in fact have two comparable qualities to both English vowels. Perhaps, the SA vowel system has only one vowel quality that maps to both English (KIT, DRESS) vowels. Since English DRESS is 100% identified and 50% of English KIT is identified as English DRESS, the SA vowel quality that maps to them would be closer to English DRESS. Such a prediction can only be confirmed by examining the vowel system of SA, which is the aim of the following chapter.

Finally, the identification of the English UC contrast (NEAR, FLEECE) shows similar results to the CG contrasts. Further analysis of the data shows that the categorised English vowel FLEECE was 100% correctly identified whereas the uncategorised vowel NEAR was 83% correctly identified, and it was 9% misidentified as English FLEECE. The results also conform to PAM predictions regarding the very good discrimination of the UC assimilation type.

A confounding finding was that 12 of the 18 English vowels examined in the identification task were 100% correctly identified by SA participants. Additionally, 17 English vowels had high percentages of correct identification, i.e. more than 80%, with the exception of English KIT vowel, which had a low percentage of correct identification (50%). As mentioned in Section 3.2.4, the stimulus material consisted of CVC words, and these words were chosen to be simple and familiar to the SA participants in order to avoid misidentification due to not knowing the lexical items. However, familiarity with the lexical items has been reported to have an effect on the identification of L2 vowels (Cebrian 2009, Flegea, et al. 1996, Mora 2005). Flege et al. (1996) tested the effect of subjective familiarity on the identification of /ɹ/ and /l/ by experienced and inexperienced Japanese learners of English compared to native English participants. Their main finding was that the Japanese learners had a greater correct identification of /ɹ/ and /l/ in familiar words, and in minimal pairs that contained familiar words than less familiar words. This finding suggests that lexical familiarity might have had an effect on the results of the present study, which led to high correct identification rates. Nevertheless, choosing familiar words was a better choice for the present study since it is testing the effect of segmental differences rather than lexical knowledge.

In sum, the identification task examined the identification of all English vowels and showed that the SA participants' discrimination and identification of the two SC contrasts was highly correct, counter to expectations. This was attributed to there in fact being enough perceptual phonetic distance between the members of these contrasts. The discrimination and the identification of the CG and UC contrasts conformed to PAM predictions, with the exception of English (KIT, DRESS) CG contrast, this was attributed to the possibility of the lack of two comparable short vowels in SA, despite the descriptions in the literature (Cowell 1964). However, the number of responses for each vowel is small and does not provide enough evidence to run statistical tests or to make any generalisations. Further examination of the perception of English vowels by SA learners will be conducted in the main study with more participants and more focused data. Additionally, the pharyngealised allophone [aːˤ] will be added to the analysis of the main study in order to determine its phonetic similarity to English BATH vowel from the SA participants' point of view, and whether this similarity had an effect on the identification of English BATH or the discrimination of the English (TRAP, BATH) vowels. Paired presentation of (FACE, SQUARE) items is needed in the main study to assess the misidentification of each vowel as the other. Also, further examination of the SA vowel system is needed, particularly to confirm the status of mid short vowels. In general, the pilot study identification task provided some answers to the research questions presented in section 3.2:

RQ1: How do SA participants perceive English vowels in terms of the English phonemic vowel inventory compared to NE participants?

The present study showed that SA participants were using their assimilation patterns from the previous task to perceive English vowels including English monophthongs and diphthongs. Nevertheless, the percentages of their identifications were highly correct for most vowels and comparable to NE participants. However, this finding should be interpreted with caution since the two SA participants had spent sometime in the UK. So, they are not the exact target learners of the present thesis.

RQ2: Which English vowel contrasts are difficult to identify or discriminate by SA participants?

The findings showed that English KIT, STRUT, LOT, GOOSE, GOAT, and NEAR were more difficult to identify. In particular, the English KIT vowel was poorly identified and discriminated from English DRESS vowel. Counter to expectations, English diphthongs such as (FACE, SQUARE) were easy to identify and discriminate. So, being a diphthong did not create a greater difficulty for SA participants to perceive. Also, English (TRAP, BATH) were unexpectedly easy to identify and discriminate.

RQ3: Is this difficulty related to quality or quantity overlap?

The difficulty that SA participants had in identifying and discriminating English (KIT, DRESS) vowels is qualitative since both are perceived as short vowels. Additionally, the misidentification is in one direction, i.e. KIT as DRESS, which suggests that the quality of their L1 /i/ vowel is closer to English DRESS. Misidentifying English NEAR as FLEECE can also be attributed to quality since both were perceived as long vowels and the misidentification is in one direction. Thus, bearing in mind that the participants' L1 is a quantitative language accounts for the difficulties they had in identifying English qualitative differences.

RQ4: To what extent can PAM assimilation categories account for the results of the present study?

PAM makes a good testable model, which accounts for some of the results of the present study. However, the SC contrasts had excellent identification and discrimination, which is unexpected by PAM. Similarly, the poor discrimination of the CG (KIT, DRESS) contrast counters what is predicted by PAM. Nevertheless, this was a pilot study with a limited number of participants, who are not the exact target learners for this thesis. Thus, these questions will be held forward to the main study for further examinations.

In the following section, we describe a production pilot experiment which was run to examine the acoustic characteristics of SA participants' L1 and L2 productions of vowels compared to NE speakers' productions.

3.3 Production Experiment

The pilot study production task was designed to examine English vowel productions of SA participants compared to NE vowel productions and their SA/L1 vowel productions. Quantitative and qualitative analyses of the results were conducted to examine the productions of a full list of English and SA vowels by SA and NE participants. However, only the results of the productions of the English vowel contrasts examined in the previous tasks and their closest equivalents in SA are presented here. A full descriptive analysis of the SA vowel system will be presented in the following chapter. Generally, the production experiment explored the following research questions:

RQ1: What are the qualitative and quantitative characteristics of English and SA vowels?

RQ2: After reducing the non-linguistic effects of sex and language using statistical normalisation procedures, which acoustic measurements reflect canonical phonetic quality and quantity differences between English and SA vowels?

RQ3: What are the acoustic characteristics of the inter-language (L2) vowel system of SA learners of English compared to their L1 (SA) and NE?

RQ4: How do the PAM perceptual predictions of the different assimilation types reflect on production patterns?

3.3.1 Hypotheses

The focus of the analysis presented here is the production of six English vowel contrasts (TRAP, BATH), (FACE, SQUARE), (KIT, DRESS), (LOT, STRUT), (GOAT, THOUGHT), and (NEAR, FLEECE) by SA participants compared to NE productions of the same vowels and their SA/L1 equivalents (/aː/), (/eː/), (/i/, /e/), (/o/, /a/), (/oː/),

and (/iː/), respectively. Based on the results of the perceptual tasks, the following hypotheses were formulated and tested acoustically:

- In the previous experiments, the English (TRAP, BATH), (FACE, SQUARE) contrasts were analysed as PAM SC contrasts, which were predicted to have poor discrimination. Nevertheless, the members of the two contrasts were highly correctly identified; and this was attributed to perceiving adequate phonetic distance between the members of the contrasts by SA participants. Since SA participants were able to identify and discriminate these vowels, they are also predicted to produce these vowels acoustically with Distinct Parameters (DP).
- As for (KIT, DRESS), (LOT, STRUT), and (GOAT, THOUGHT), which were analysed as PAM CG contrasts, they were predicted to have from moderate to good discrimination. English (LOT, STRUT) and (GOAT, THOUGHT) showed good identification and discrimination, whereas English (KIT, DRESS) had poor discrimination, and low identification of the KIT vowel. In production, English (LOT, STRUT), and (GOAT, THOUGHT) are predicted to be produced acoustically distinct from each other (DP), whereas the members of English (KIT, DRESS) contrast are predicted to be produced with Overlapped acoustic Parameters (OP).
- English (NEAR, FLEECE) contrast was analysed as PAM UC assimilation type, and as predicted the members of this contrast had very good identification and discrimination. In production, these are also predicted to be produced acoustically distinct (DP).

3.3.2 Participants

The same two SA and NE participants who participated in the previous pilot study experiments also took part in the production task. As mentioned earlier, the participants were asked to rate their language skills on a seven point scale. The rate for their speaking skill (though high) was the least (6/7) and this suggested that SA speakers would have greater difficulty producing what matches the NE productions, for more information about the participants see Section 3.1.3 and Section 3.2.3.

3.3.3 Recordings

SA and NE participants were recorded following the same procedures used to record NE speakers in the previous experiments, see Section 3.1.2. However for this production task, the English stimuli and the instructions were the same for both SA and NE speakers.

3.3.4 Stimulus Material

3.3.4.1 SSBE Stimuli

SA and NE speakers produced five randomised blocks of keywords containing 11 English monophthongs and 7 diphthongs, which are presented in Table 3-8 below. However, the main interest of the present study was to examine the productions of the members of the six English vowel contrasts, which were found to be confusing for SA learners in their categorisations and/or identifications. These vowel contrasts are: (TRAP, BATH), (FACE, SQUARE), (KIT, DRESS), (LOT, STRUT), (GOAT, THOUGHT), and (NEAR, FLEECE).

Table 3-8: English vowels and the target words produced by NE and SA participants in the production task

Monophthongs				Diphthongs			
SSBE vowel	Target Word	Phonemic transcription	Similar to target	SSBE vowel	Target word	phonemic transcription	Similar to target
FLEECE	heed	/hiːd/	feed	FACE	hayed	/heɪd/	hate
KIT	hid	/hɪd/	sit	GOAT	hoed	/həʊd/	goat
DRESS	head	/hɛd/	red	NICE	hide	/haɪd/	nice
TRAP	had	/hæd/	had	MOUTH	howdy	/haʊdɪ/	house
STRUT	hudd	/hʌd/	cut	CHOICE	hoyed	/hɔɪd/	noise
BATH	hard	/hɑːd/	hard	NEAR	hear	/hɪə/	near
LOT	hod	/hɒd/	hot	SQUARE	hair	/hɛə/	hair
THOUGHT	hawed	/hɔːd/	thought				
FOOT	hood	/hʊd/	book				
GOOSE	who'd	/huːd/	food				
NURSE	heard	/hɜːd/	hurt				

All target vowels were put in the monosyllabic /hVd/ context (except the words ending with <r>), as presented in Table 3-8 above. The English target words were put in the phrase "Say _____ again". After examining vowels in several contexts Stevens and House (1963: 116) found that "/hVd/ context has a negligible effect upon the articulation during the central portion of the vowel, that is, the vowel in the /hVd/ context is generated with essentially the same articulatory configuration as the vowel in isolation". Other researchers such as Yang (1996: 248) used the /hVd/ context because the vowel in this context "should not exhibit coarticulatory effects of the preceding consonant because /h/ is the voiceless variant of the following vowel; alveolar /d/ also has relatively little influence on the formants of the preceding vowel". Thus, this context could be used reliably to get vowel measurements close to those in isolation. Table 3-8 presents the phonemic rather than the phonetic transcription of the target words. A canonical pronunciation of each vowel was required, because this is what could be considered as the English input for the SA participants, which might be different from the actual pronunciation of NE speakers. For example, a NE teacher commented that it was difficult to teach Arab students the unstressed form /ə/ and that those students considered it as inappropriate pronunciation (personal communication).

The best three productions of each vowel, which had no noise, hesitation, or mispronunciation, were extracted for each participant and analysed acoustically. The analysis here covers the productions of the six English vocalic contrasts mentioned earlier. All together, 3 repetitions × 12 vowels × 4 speakers = 144 items were analyzed.

3.3.4.2 Syrian Arabic Stimuli

In addition to the SA participants' English (L2 henceforth) productions of the English vowels, SA participants were asked to produce comparable SA (L1 henceforth) stimuli. Therefore, SA participants' L2 productions were compared to both their L1 and NE productions. SA participants were asked to produce five randomized blocks of the five SA (L1) long vowels (/iː/, /eː/, /aː/, /oː/, /uː/), five short vowels (/ɪ/, /e/, /a/, /o/, /u/), schwa /ə/, and vowel-glide sequences /aj/ and /aw/.

All SA target vowels were put in the same monosyllabic /hVd/ context used for the English stimuli, as illustrated in Table 3-9 and Table 3-10 below. The target words were put in the phrase /ktoːb ___ marteːn/ “Write ___ twice”.

Table 3-9: SA (L1) long vowels and the target /hVd/words used in the production task

Long vowels				
SA vowel	Target word	Similar to target	Arabic word	English gloss
/iː/	/hiːd/	/fiːd/	فيد	do something useful!
/eː/	/heːd/	/zeːd/	زيد	proper name
/aː/	/haːd/	/haːd/	هاد	this one
/oː/	/hoːd/	/xoːd/	خود	Take!
/uː/	/huːd/	/huːd/	هود	prophet name

Table 3-10: SA (L1) short vowels and the target /hVd/ words used in the production task

Short vowels				
SA vowel	Target word	Similar to target	Arabic word	English gloss
/i/	/hɪd/	/hɪdd/	هِد	destroy!
/e/	/hed/	/naːhed/	ناهد	proper name
/a/	/had/	/hadd/	هَد	he destroyed
/o/	/hod/	/hidhod/	هدهد	kind of bird
/u/	/hud/	/huda/	هدى	proper name
/ə/	/hədne/	/hədne/	هدنة	a truce

Since it was difficult to have real /hVd/ words in SA for all target vowels, some nonsense words were used. To ensure that the SA speakers produce the target vowels correctly, a real monosyllabic /CVC/ word, that had the same target vowel as the one in the nonsense /hVd/ context, was presented simultaneously alongside the nonsense word. The speakers were asked to produce the target nonsense word with the same vowel as in the real word beside it.

According to Cowell's (1964) description of the Damascene sound system, there are restrictions on the occurrences of the six short vowels of SA. For example, /ə/ does not occur word finally, /e/ and /o/ never occur accented and rarely in open syllables except word-finally, and /i/ and /u/ do not occur before word-final consonants. Since it was difficult to find monosyllabic words with short vowels, di-syllabic words were chosen for the model real words, see Table 3-10 above. For the vowel-glide sequences, SA words /fay/ ‘shadow’ and /dˤaw/ ‘light’ were used.

Similar to the English stimuli, the best three productions of each vowel for each participant were extracted and analysed. Here, the effect of L1 is tested on L2 productions of the six English vocalic contrasts under investigation. Thus, only (/aː/, /eː/, /ɪ/, /e/, /o/, /a/, /oː/, /iː/) SA vowels, which were classified to be similar or close to the English vowels under investigation, were analysed. Together the SA stimuli included 3 repetitions × 8 vowels × 2 speakers = 48 items.

3.3.5 Procedures

The English and the SA stimulus material were presented to the participants using a PowerPoint slide-show, and each slide contained one phrase. The phrases were randomized and then presented to the participants. This was a self-paced task: the speakers were asked to read each phrase in a natural way and in a normal speed. When they finished, they were directed to proceed to the next phrase by clicking the arrows on the keyboard.

The English stimulus material was presented in conventional English orthography, and the Arabic stimulus material was presented in conventional Standard Arabic orthography. Nevertheless, the participants were asked to read the Arabic phrases in their local SA and not in standard Arabic. It can be argued that asking Arabic speakers, in general, to read with a local variety of Arabic is not as difficult as it used to be, particularly for young people who have been recently using Standard Arabic written form to represent dialectal Arabic in their personal daily communication. For example, Arabic speakers are using Standard Arabic written form to represent their dialect in their texting and emails.

However, caution was taken when presenting the written Arabic data. Although Arabic spelling is virtually phonetic, only long vowels are represented in writing whereas the short vowels are usually represented as diacritics (Smith 2001). The diacritics (short vowels) are not usually added unless required in careful writing. To make sure that the SA participants produced the target words correctly, firstly, they were acquainted with the stimulus material before starting with the recordings. Secondly, when having difficulty with a particular word, the participants were given a familiar word with the same vowel. The latter procedure was also applied to elicit the English stimulus material. SA speakers were asked to read the English stimuli before the SA ones, as they were expecting an English task and this helped to keep them oriented to English mode.

3.3.6 Data Analysis

The best three productions of each vowel, with no hesitation or noise, were extracted and analysed acoustically for each subject using Praat 5.1.14 (Boersma and Weenink 2009). The analysis procedures were automated using a Praat script. The acoustic landmarks being investigated were identified as time points using TextGrid files. Spectrograms were used for segmentation and waveforms for more fine-grained segmentation decisions (Yang 1996). I placed the boundaries carefully by hand using waveforms and wide-band spectrograms in addition to auditory verification.

For this study, the first two formants F1-F2 and vowel duration were extracted. The first two formants were extracted using formant tracks which were generated using the Burg algorithm in Praat. All tracks were generated using 0.025 ms window length, 50 Hz pre-emphasis, and 5500 Hz spectrogram view range. The beginning of the vowel was identified manually as the end of noise for /h/ and beginning of the periodic waveform, whereas the end of the vowel was identified as the end of the periodic waveform and the beginning of the closure period of /d/, as can be seen in Figure 3-5 below (Paolo, et al. 2011: 91, Yang 1996). Following McDougall and Nolan (2007), frequency measurements were extracted from +10% automatically selected steps (as illustrated in Figure 3-5), which was useful to show formant dynamics and trajectories. This was particularly helpful to visualize monophthongs and diphthongs throughout the vowel. However, for the purpose of this study, only the midpoint frequency measurement, which represents the steady state of the vowel in most cases, was examined. In Figure 3-5 below, point 5 represents the midpoint of the vowel. The figure also shows vowel duration, which was measured by subtracting the time of the ending point of the vowel (EV) from the time of the starting point of the vowel (SV).

Additionally, the analysis of the production of the three English vowel contrasts which included diphthongs, (FACE, SQUARE), (GOAT, THOUGHT), and (FLEECE, NEAR) and their assimilated long categories in SA (/eː/, /oː/, and /iː/), included extracting frequency measurements from three points: (25%, 50%, and 75%) throughout the vowel, in order to track diphthongal trajectories. Thus, for these vowels 25% and 75% frequency measurements were examined in addition to midpoint frequencies.

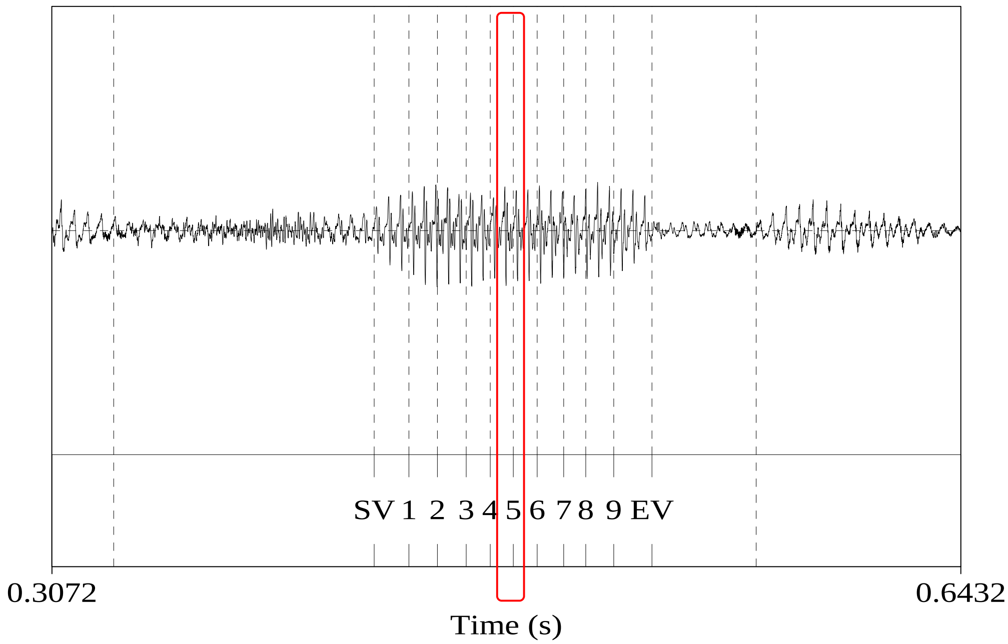

Figure 3-5: a wave form and text grid of the word [hard] as produced by a female English speaker (E3), it shows the beginning (SV) and ending (EV) of the vowel, the +10% points inserted, and the midpoint (5) where frequency measurements were extracted.

Since all measurements were extracted automatically using a Praat script, the measurements were checked and explored using SPSS to spot outliers. For example, Praat might read close F1 and F2 frequency values for vowels like GOOSE and GOAT both as the first formant F1; and it will read F3 as F2 in this case (Paolo, et al. 2011: 94). Any outliers were re-examined manually in order to verify their values.

In the following section, the results and findings are presented for each English contrast separately. Only descriptive analysis of the results is presented here because the number of items representing each vowel is small: 3 repetitions * 2 participants, which mean that there are 6 tokens for each vowel. Furthermore, the acoustic measurements for the pilot study were not normalised. However, the findings cannot be conclusive without normalising the data across speakers and sex. Therefore, the acoustic measurements should be normalised in the main study in Chapter 7.

3.3.6.1 (TRAP, BATH): (/æ/, /ɑː/)

The production of the English (TRAP, BATH) contrast was compared to their equivalent vowel in SA /aː/. The acoustic analysis of these vowels included comparing vowel durations and F1 and F2 frequency measurements. Figure 3-6 presents mean vowel duration of English (TRAP, BATH) as produced by NE and SA participants (L2), and mean duration of SA /aː/ as produced by the same SA participants (L1). As can be seen, the duration difference between NE productions of the English TRAP and BATH vowels is greater than L2 productions of the same vowels. Compared to English productions, the duration of SA /aː/ vowel is close to the duration of English TRAP. As a possible result of an L1 effect, SA participants seem to produce both English vowels with a duration that is close to the SA /aː/ vowel.

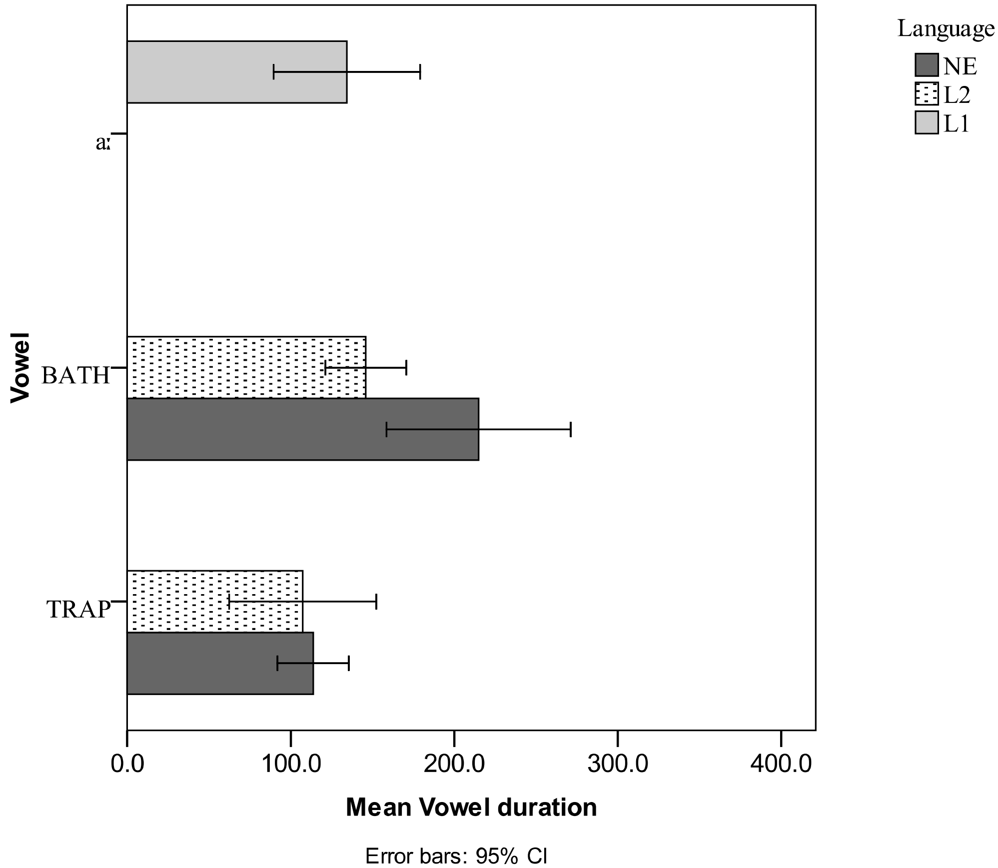

Figure 3-6: mean vowel duration of English TRAP and BATH (/æ/-/ɑː/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowel in L1 /aː/

As for frequency measurements, mean midpoint F1 and F2 frequencies for NE and L2 productions compared to their L1 equivalent are presented in Figure 3-7 and Figure 3-8. Figure 3-7 shows that mean F1 and F2 frequencies of NE and L2 productions of TRAP compared to SA /aː/ are very close to each other in the vowel space. Figure 3-8, on the other hand, shows that NE production of BATH is higher and backer than both L2 and L1 productions. Thus, the SA /aː/ vowel is closer to English TRAP in the vowel space as well. Additionally, L2 productions of English TRAP and BATH vowels are acoustically distinct from each other, but not in the same way as NE productions. Thus, the results support the hypothesis for the L2 production of this contrast, i.e. as predicted the SA participants produced English (TRAP, BATH) vowels with Distinct Parameters (DP).

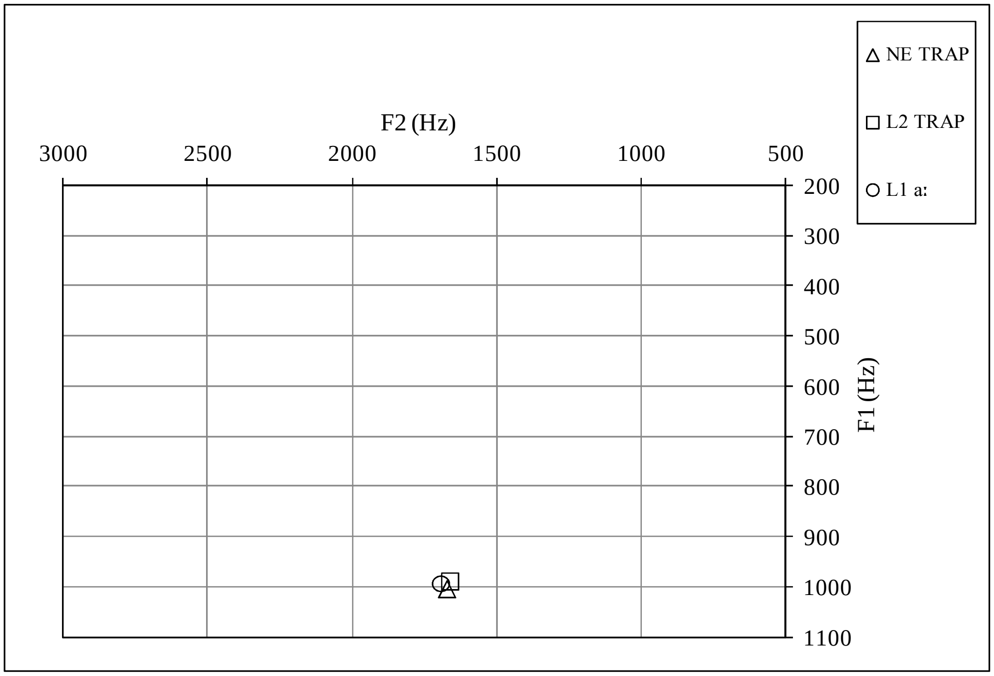

Figure 3-7: Mean midpoint F1 and F2 frequencies of English TRAP /æ/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /a:/

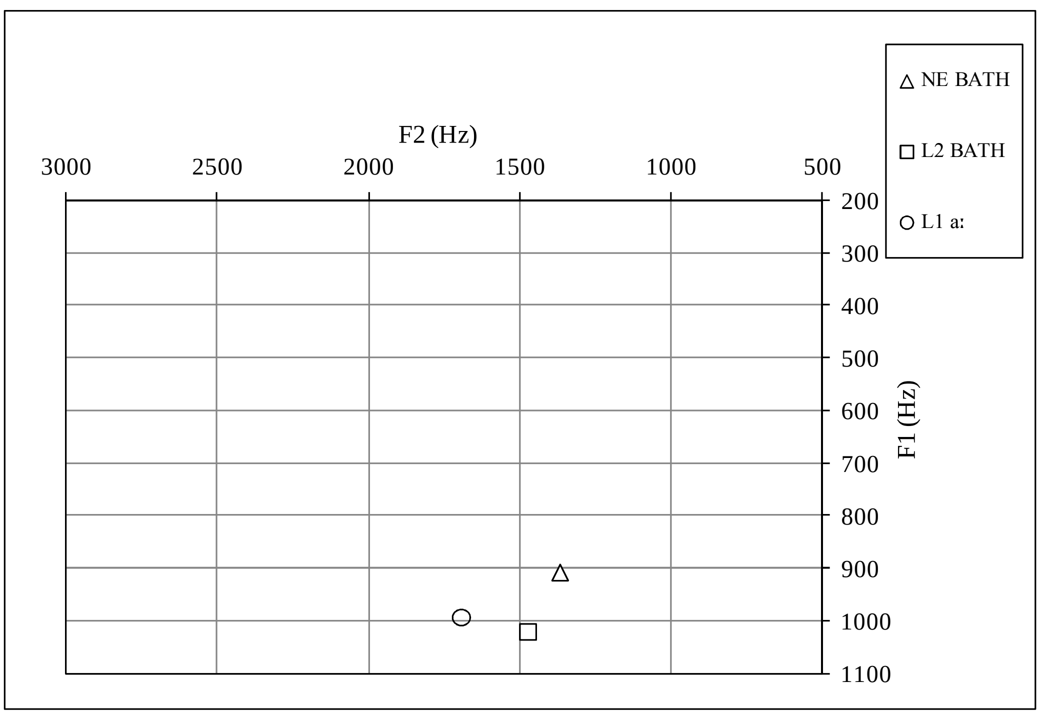

Figure 3-8: Mean midpoint F1 and F2 frequencies of English BATH /ɑ:/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /a:/

3.3.6.2 (FACE, SQUARE): (/eɪ/, /ɛə/)

Similar to the previous contrast, the English diphthongal FACE and SQUARE (/eɪ/-/ɛə/) contrast was produced with a greater durational difference by NE participants than L2 participants. Figure 3-9 shows that NE productions are longer in duration than L2 productions of the same vowels, and L1 /eː/ production is shorter than both NE and L2 productions.

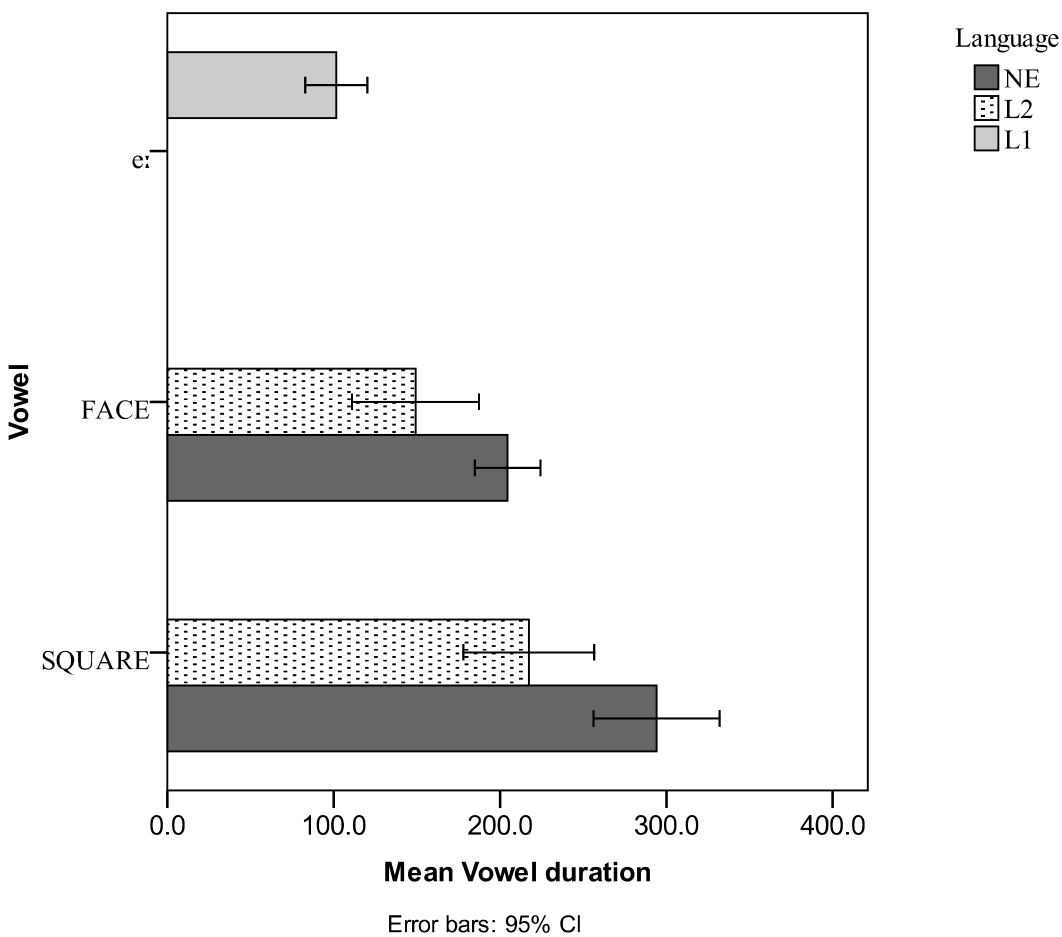

Figure 3-9: mean vowel duration of English FACE and SQUARE (/eɪ/-/ɛə/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowel in L1 /eː/

Figure 3-10 below shows that NE production of FACE vowel has a clear diphthongal movement in the vowel space from a central mid-high into a front-high vowel quality. L2 productions of the same vowel show a slight movement in the vowel space towards a fronter vowel quality. As predicted, the monophthongal L1 production of SA /eː/ vowel does not show any movement in the vowel space. Figure 3-11 also shows that NE productions of the SQUARE vowel have a very clear movement in the vowel space from a central mid-low into a front mid vowel quality, whereas L2 productions of the same vowel show a very small movement. L2 productions of the English (FACE, SQUARE) contrast are greatly affected by L1 production of SA mid-long vowel /eː/, which is very different from both English vowels in terms of quantity and quality. L2 productions of the two vowels are not very distinct from each other, and both are produced close to L1 /eː/. Thus, counter to expectations L2 productions of this contrast were not produced with DP, they were produced with Overlapping parameters (OP). This finding suggests that the participants' assimilation patterns from the PAT were reflected in their production patterns, which can be argued to be a greater L1 effect in both tasks than on the identification task, in which these vowels were highly correctly identified and discriminated. The implications of the L1 effect on the three tasks will be elaborated on in the main study Chapters 5-7.

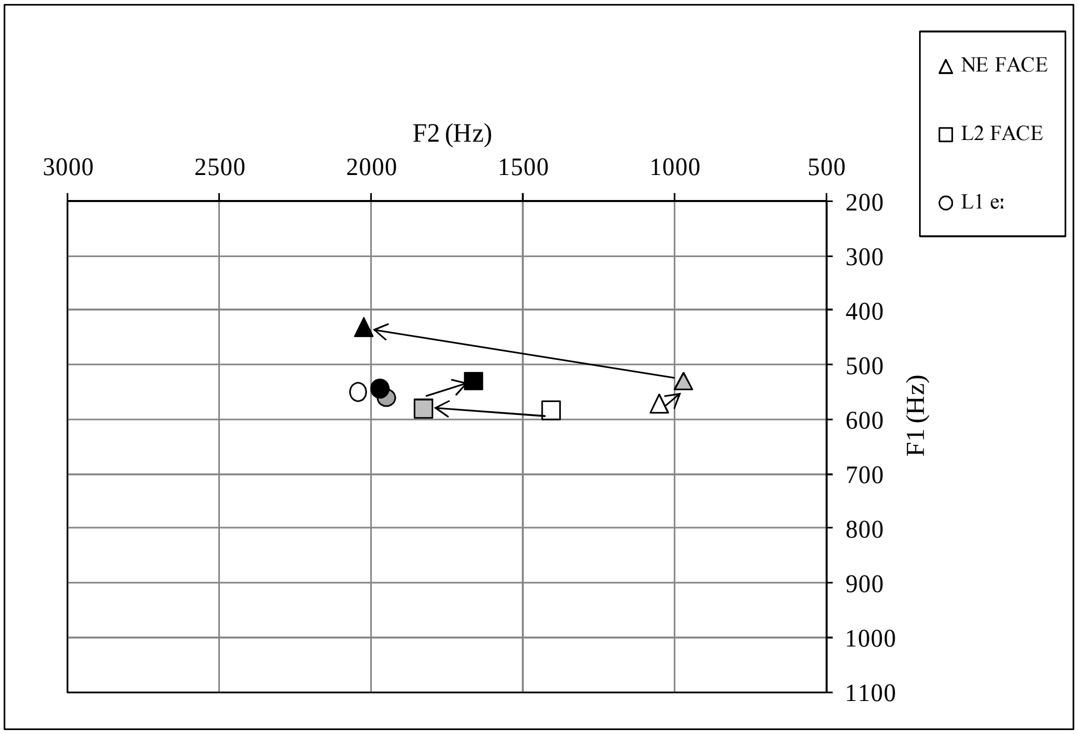

Figure 3-10: mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE FACE /eɪ/ productions compared to L2 productions of the same vowel and L1 /eː/ productions.

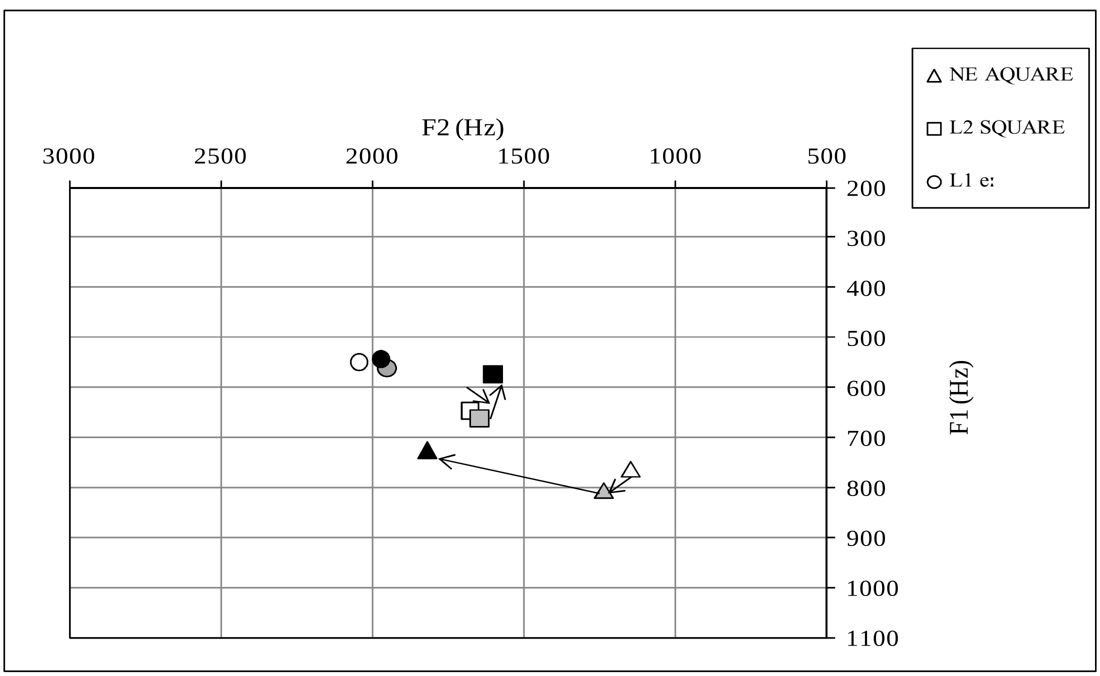

Figure 3-11: mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE SQUARE /ɛə/ productions compared to L2 productions of the same vowel and L1 /eː/ productions.

3.3.6.3 (KIT, DRESS): (/ɪ/, /ɛ/)

NE productions of KIT and DRESS were compared to L2 productions of the same vowels and to their equivalents in SA/L1 /i/ and /e/, respectively. Figure 3-12 shows that the duration of L2 productions are longer than NE productions. However, the durational difference between the two English vowels is the same for NE and L2 productions. As can be seen, SA/L1 vowels are shorter in duration than both NE and L2 productions; and the durational difference is also less for SA (/i/, /e/) vowels.

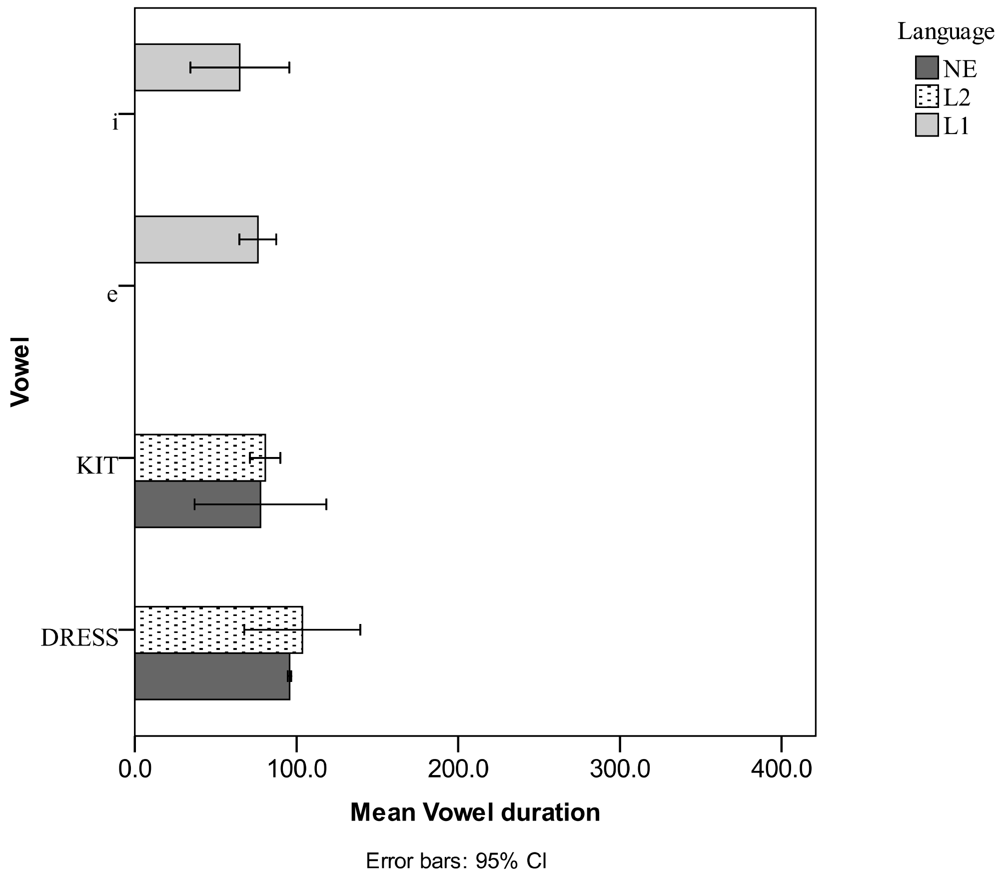

Figure 3-12: mean vowel duration of English KIT and DRESS (/ɪ/-/ɛ/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowels in L1 /i/ and /e/

As for the frequency measurements, Figure 3-13 shows that the NE KIT vowel is lower and fronter than both L2 and L1 productions, which are closer to each other in height and place. Figure 3-14 also shows that the NE DRESS vowel is lower than both L2 and L1 productions, which are almost identical. Further examination of L1 /i/ and /e/ vowels shows that they are very close to each other in the vowel space, which suggests that they might represent one category that is closer to the NE KIT vowel. However, based on two speakers' productions only we cannot confirm if SA has one or two short front vowel categories. The figures also show that L2 productions of the English KIT and DRESS vowels are distinct from each other in height but not in place. Thus, counter to expectations, the L2 productions of this contrast were produced with overlapping place, but distinct duration and height. However, since no statistical analysis was carried out, the production of this contrast by SA participants is argued to be with DP since there were differences in terms of height and duration, and this is counter to what was predicted.

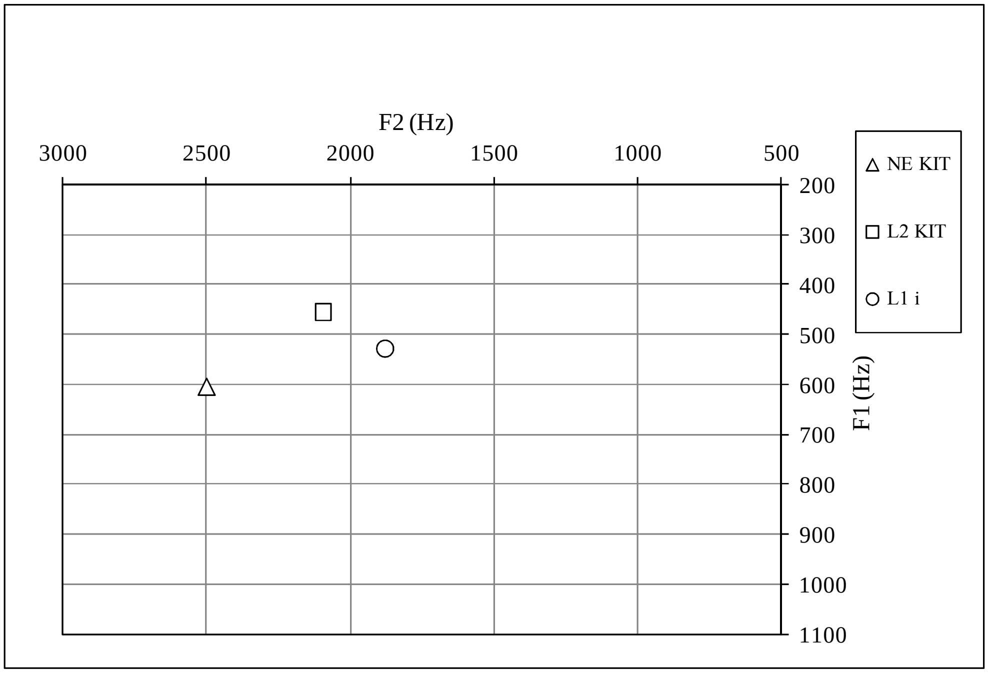

Figure 3-13: Mean midpoint F1 and F2 frequencies of the English KIT /ɪ/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /i/

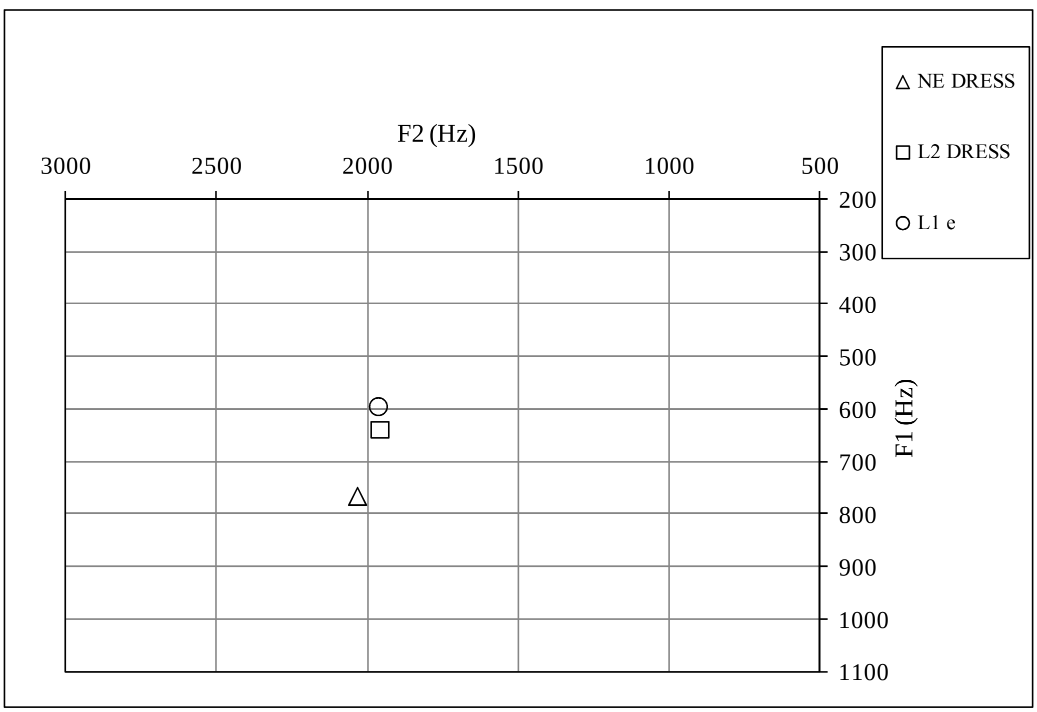

Figure 3-14: Mean midpoint F1 and F2 frequencies of the English DRESS /ɛ/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /e/

3.3.6.4 (LOT, STRUT): (/ɒ/, /ʌ/)

The productions of the NE LOT and STRUT vowels are compared to L2 productions of the same vowels and their SA/L1 equivalents /o/ and /a/, respectively. The vowel durations of L2 productions were very close to NE productions which were longer than SA/L1 vowel durations, as can be seen in Figure 3-15.

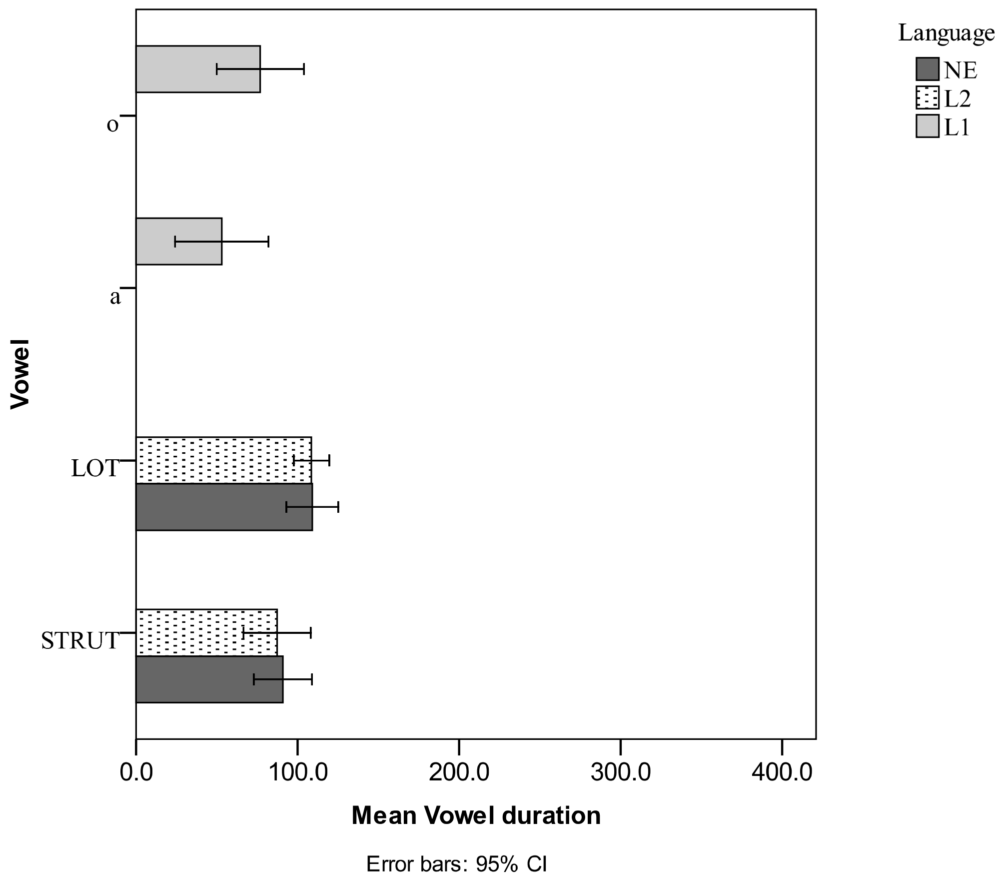

Figure 3-15: mean vowel duration of English STRUT and LOT (/ʌ/, /ɒ/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowels in L1 /a/ and /o/

Figure 3-16 below shows that the production of the NE and L2 LOT vowel are close to each other and both have lower and backer quality than SA/L1 vowel /o/. The same applies to the production of NE and L2 STRUT vowel (Figure 3-17); both are lower and slightly backer than SA/L1 vowel /a/. However, L2 STRUT is much lower than the NE production of the same vowel. As can be seen in the figures, the NE (LOT, STRUT) vowels differ mainly in place, where LOT is back and STRUT vowel is central. Also, SA (/o/, /a/) vowels differ mainly in place, where /a/ is central and o is back. Unlike NE and L1 contrasts, L2 productions of LOT and STRUT vowels differ mainly in height, where LOT is higher in the vowel space. Therefore, this finding suggests that L2 productions were produced with DP, though they differed from those produced by the NE participants.

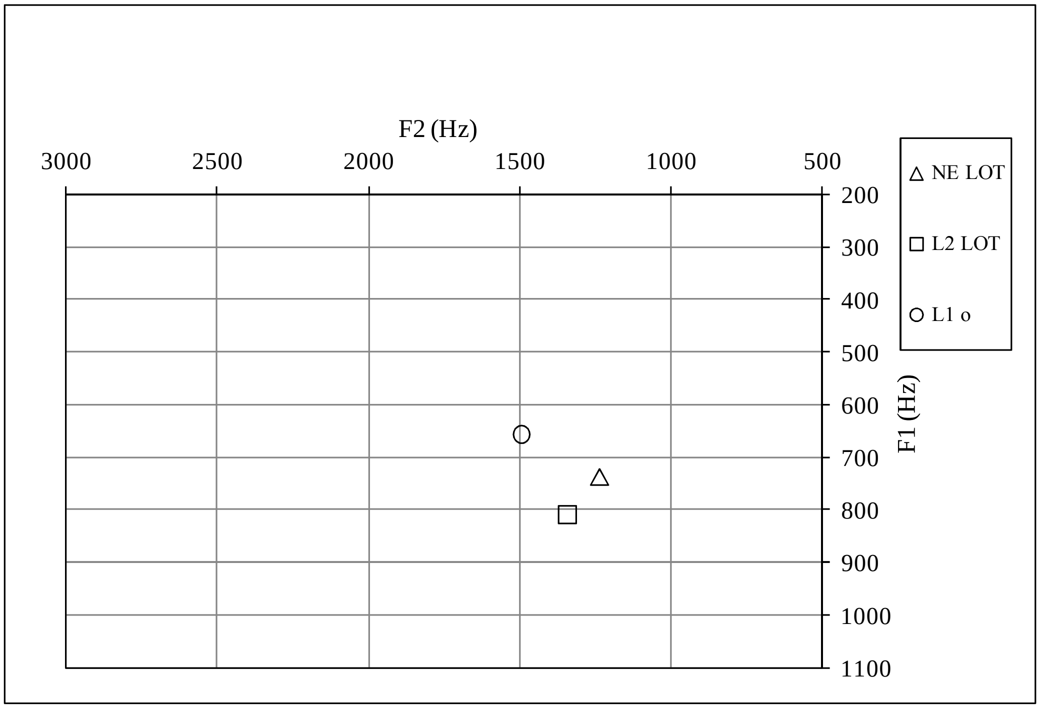

Figure 3-16: Mean midpoint F1 and F2 frequencies of the English LOT /ɒ/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /o/

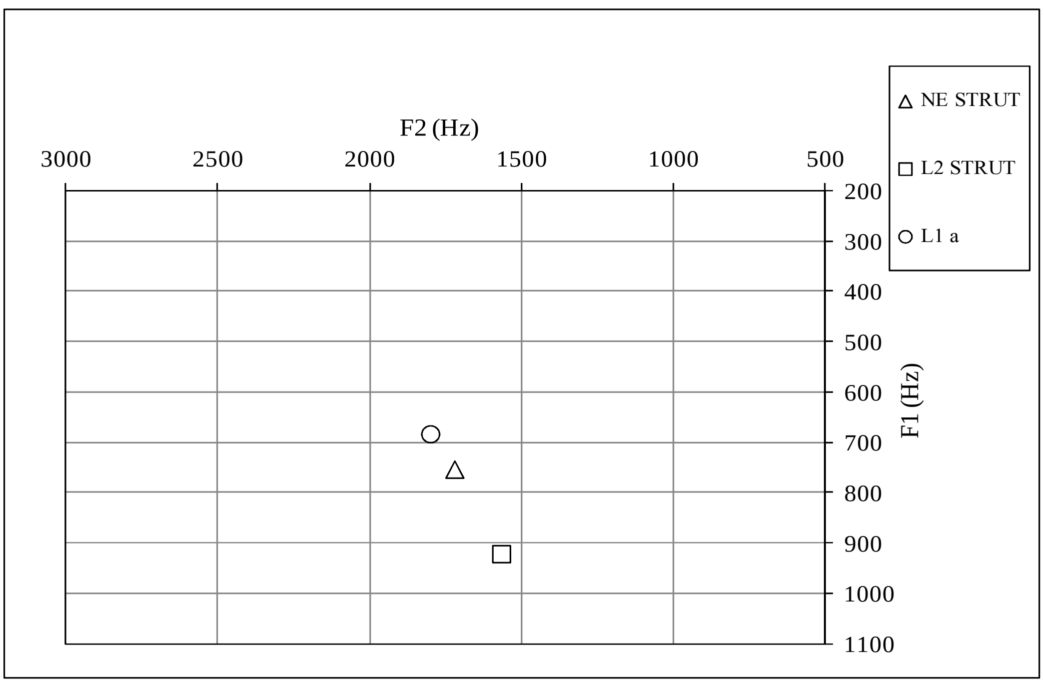

Figure 3-17: Mean midpoint F1 and F2 frequencies of the English LOT /ʌ/ vowel as produced by NE and L2 participants, and mean F1 and F2 frequencies of their equivalent vowel in SA/L1 /a/

3.3.6.5 (GOAT- THOUGHT): (/əʊ/, /ɔː/)

Figure 3-18 below presents mean vowel duration of the NE GOAT and THOUGHT vowels compared to L2 productions of the same vowels and SA/L1vowel /oː/. As can be seen, NE productions have longer durations than both L2 and L1 productions. Counter to expectations, the L2 production of THOUGHT vowel is much shorter than NE THOUGHT production, i.e. less than 50%. A possible explanation is that the L2 THOUGHT vowel here is produced with an American accent, in which the THOUGHT vowel is produced shorter and more open than British English (Mcmahon 2002: 81). However, the L2 productions have a greater durational difference between the two English GOAT and THOUGHT vowels than NE productions. L1 vowel duration is shorter than both English vowels, and it is different from the SA participants' L2 vowel productions.

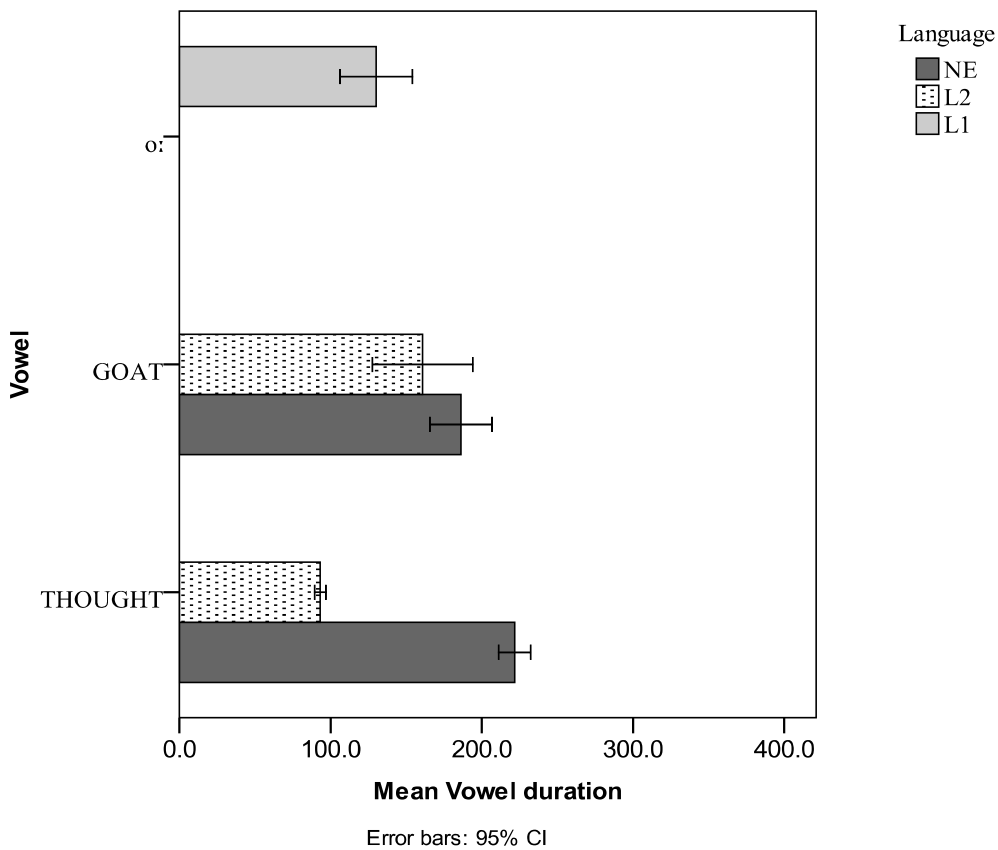

Figure 3-18: mean vowel duration of the English HOED and HAWED (/əʊ/- /ɔː/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowel in L1/oː

In the vowel space, L2 productions of English GOAT vowel are different from NE productions. Figure 3-19 below shows that L2 productions are backer and do not show trajectory movement as much as NE productions do. L2 productions of GOAT are closer to SA/L1 productions of /oː/ vowel. Figure 3-20, on the other hand, shows that L2 production of the English THOUGHT vowel has a trajectory movement, which is not produced by NE participants. It also shows that NE THOUGHT and L1 /oː/ productions are in the same area in the vowel space whereas L2 production is lower than both NE and L1 productions. This latter finding supports the earlier suggestion that the L2 THOUGHT vowel was produced with an American accent rather than British. Thus, the SA /oː/ vowel is similar to English THOUGHT vowel, and SA participants produce the GOAT vowel with a quality that is also similar to THOUGHT (might be a case of hypercorrection of to which vowel they need to assign a trajectory movement). As can be seen from the figures, L2 learners managed to produce these English vowels with distinct parameters, however, those parameters did not exactly match NE productions.

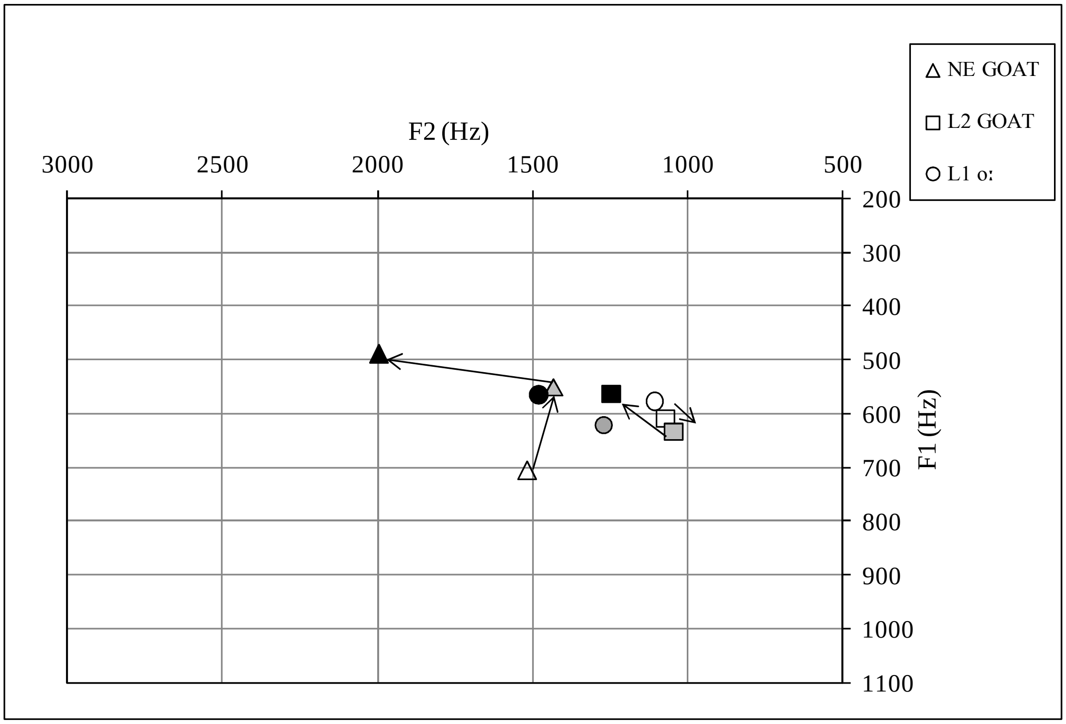

Figure 3-19: : mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE GOAT /əʊ/ productions compared to L2 productions of the same vowel and L1 /oː/ productions.

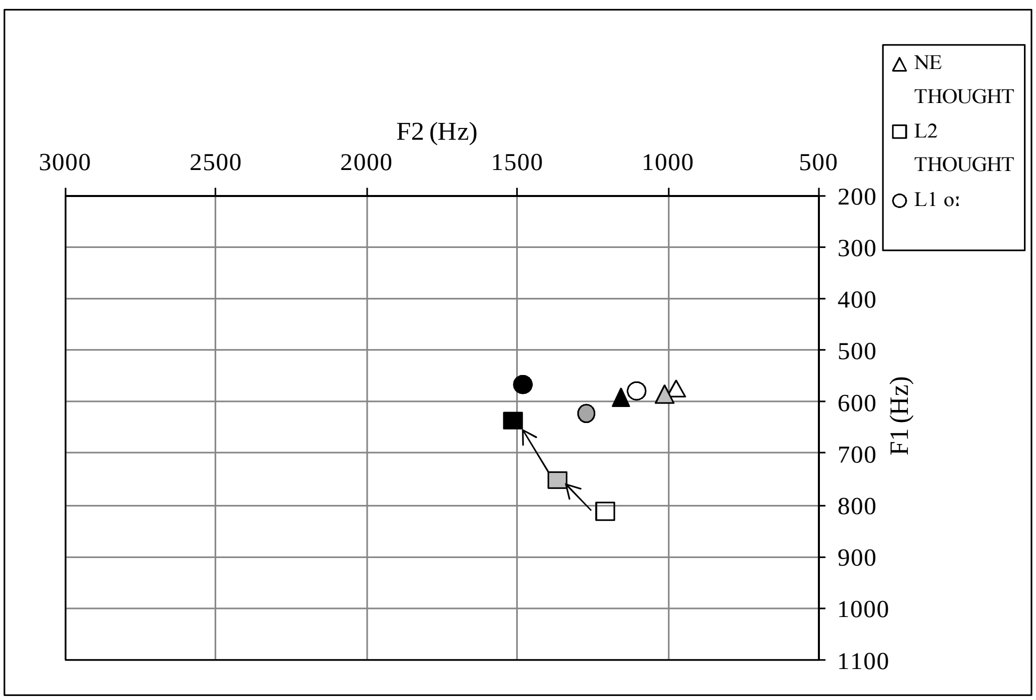

Figure 3-20: mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE THOUGHT /ɔː/ productions compared to L2 productions of the same vowel and L1 /oː/ productions.

3.3.6.6 (NEAR, FLEECE): (/ɪə/, /iː/)

Figure 3-21 shows the mean vowel duration of L2 NEAR and FLEECE (/ɪə/-/iː/) productions compared to NE and L1 /iː/ productions. The results show that NE productions had longer duration than L2 productions of the same vowels and the L1 /iː/ vowel. However, L2 productions have a greater duration difference between the two English vowels compared to NE productions. The duration of L2 FLEECE and SA/L1 /iː/ vowel is almost the same which means that SA participants produced the English FLEECE vowel with the duration of their L1 /iː/ vowel.

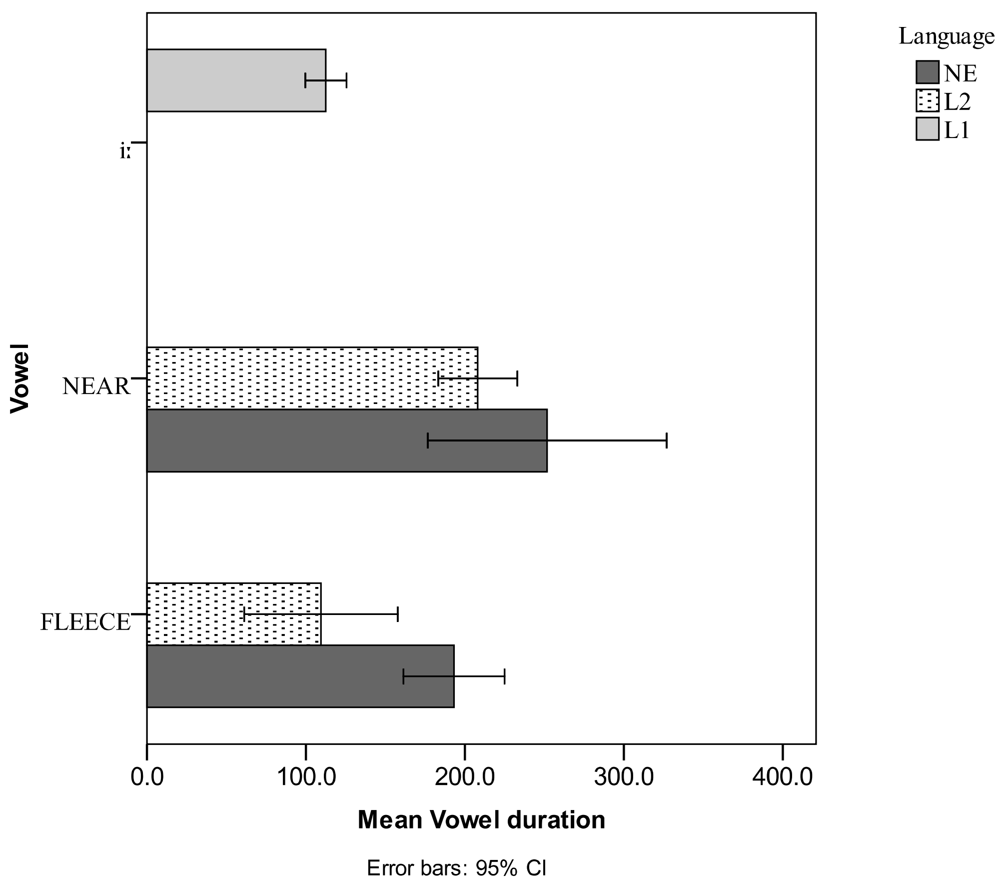

Figure 3-21: mean vowel duration of the English NEAR and FLEECE (/ɪə/-/iː/) vowels as produced by NE, L2 participants, and mean durations of their equivalent vowel in L1/iː/

The results of the frequency measurements show that L2 productions of the FLEECE vowel are half way between NE and L1 productions, which had similar height, but NE productions were fronter in quality, as shown in Figure 3-22. On the other hand, Figure 3-23 shows that L2 productions of the NEAR vowel are half way between NE and L1 productions, which are similar in place but not in height; the NE NEAR vowel is much lower than the L1 /iː/ vowel. L2 productions of the NEAR vowel show a small trajectory movement, but it does not match NE productions. Thus, SA participants produced the two English vowels distinct from each other (DP). However, their productions do not match NE or L1 productions.

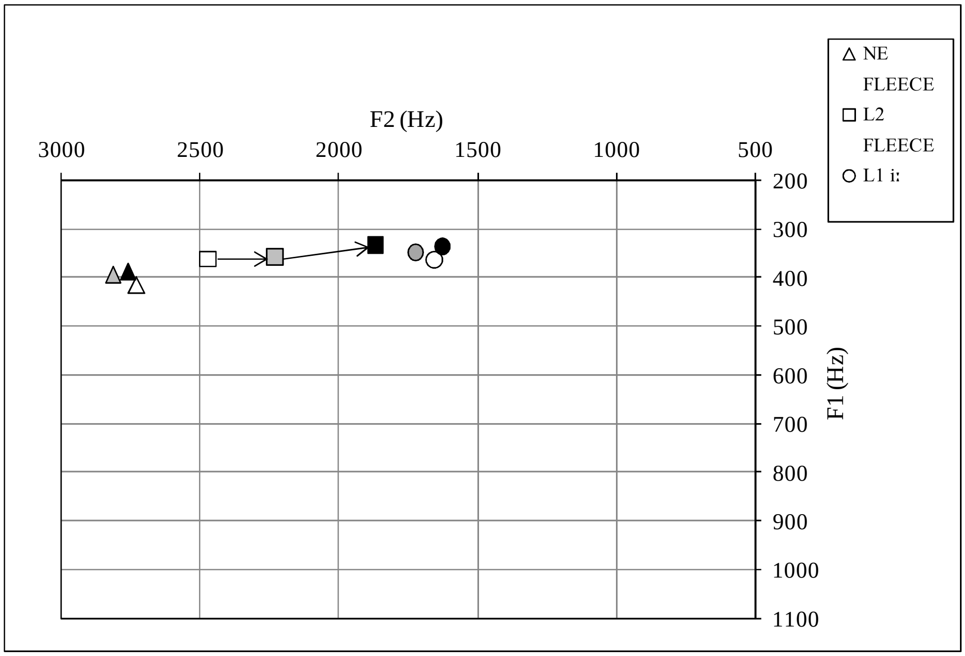

Figure 3-22: mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE FLEECE /iː/ productions compared to L2 productions of the same vowel and L1 /iː/ productions.

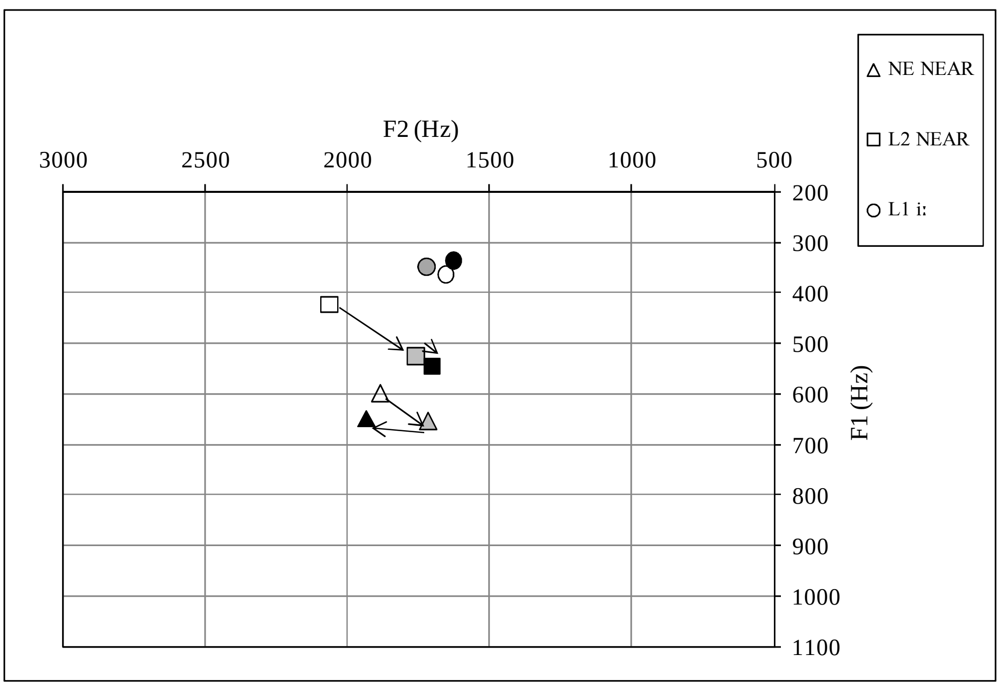

Figure 3-23: mean F1 and F2 frequencies of 25% (unfilled), 50% (grey), and 75% (black) of NE NEAR /ɪə/ productions compared to L2 productions of the same vowel and L1 /iː/ productions.

In sum, the production task managed to provide some answers to the following research questions, with the focus on the above English contrasts:

RQ1: What are the qualitative and quantitative characteristics of English and SA vowels?

The pilot study provided acoustic analysis and description of L2 productions of six English vowel contrasts compared to NE productions and to their SA equivalent vowels.

RQ2: After reducing the non-linguistic effects of sex and language using statistical normalisation procedures, which acoustic measurements reflect canonical phonetic quality and quantity differences between English and SA vowels?

The results showed that English and SA equivalent vowels, such as FLEECE and SA /iː/, were found to differ in terms of quantity and quantity. Since the acoustic measurements were not normalised, the answer to this question is speculative and not conclusive. Therefore, the measurements will be normalised in the main study to be able to answer this question.

RQ3: What are the acoustic characteristics of the inter-language (L2) vowel system of SA learners of English compared to their L1 (SA) and NE?

L2 productions of the SA participants were mostly affected by their L1 vowels, particularly equivalent vowels such as TRAP and FLEECE. In terms of duration, NE productions tended to be longer than L2 and L1 productions. As for quality, L2 productions tended to be half way between their L1 and NE productions.

RQ4: How do PAM perceptual predictions of the different assimilation types reflect on production patterns?

The assimilation patterns of the PAM were used to predict the identification and discrimination patterns of English vowel contrasts. The results of the latter task were then used to predict the production patterns of the same vowels. As shown throughout this chapter, the assimilation patterns do not explain all the identification and discrimination patterns, which do not also explain all the production patterns. The answer to this questions would be that it is likely that the participants were using or referring to different processes when they were asked to assimilate, identify, or produce the target vowels, which would explain the mismatch between the findings of the three experiments. PAM makes a good model to build up identification and production predictions; however, it needs to take into consideration the differences between the processes (mental/articulatory, memory load, consciousness) associated with each task.

3.4 Summary

A summary of the results of the three experiments of the pilot study, i.e. the PAT, the identification task (which included identification and discrimination results), and the production task, is given in Table 3-11 below. The results of the PAT were used to predict the L2 identification and discrimination patterns, and then the results of the latter were used to predict the production patterns of the same vowels (DP or OP). The table shows that the English SC contrasts had excellent identification and discrimination, which contradicts the PAM prediction. L2 productions of English SC (TRAP, BATH) vowels were acoustically distinct from each other and close to NE productions. The perceptual and acoustic results of L2 English (TRAP, BATH) suggest that this is a TC (Two Category) contrast rather than a SC contrast, in which TRAP is categorised as SA /aː/ and BATH is, arguably, categorised as SA allophonic [aːˤ] which occurs after pharyngealised segments in Arabic. This latter argument accounts for the excellent identification and discrimination of English (TRAP, BATH) vowels by SA participants. Thus, SA allophone [aːˤ] will be added for further examination in the main study.

Table 3-11: summary of the PAT, identification task, and production experiment results of the pilot study

English contrast	PAM classification	Identification	Discrimination	Production
(TRAP, BATH)	SC/ reanalysed as TC	(excellent, excellent)	excellent	DP
(FACE, SQUARE)	SC	(excellent, excellent)	excellent	OP
(KIT, DRESS)	CG	(poor, excellent)	poor	DP
(LOT, STRUT)	CG	(very good, very good)	very good	DP
(GOAT, THOUGHT)	CG	(very good, excellent)	excellent	DP
(FLEECE, NEAR)	UC	(excellent, good)	very good	DP

On the other hand, the perceptual and acoustic analysis of the results showed that English (FACE, SQUARE) vowel contrast is a SC contrast. Furthermore, the excellent identification and discrimination of the two vowels by SA participants was attributed to perceiving enough phonetic distance between them. However, L2 productions of the English (FACE, SQUARE) vowels were not acoustically very distinct from each other in the vowel space; they were produced with overlapping acoustic parameters (OP). Since the two members of this contrast are diphthongs and diphthongisation does not have any phonological role in SA, the participants were not able to produce the two vowels distinct from each other and the production of both vowels was not diphthongal in nature. Also, both productions were closer to SA vowel monophthongal mid long vowel /eː/. Thus, SA participants were able to perceive a phonetic contrast which they could not produce, which coincides with the general assumption that perception precedes production (Best 1999). Additionally, L2 exposure of the SA participants of the present study was limited and was mostly classroom based. In order to produce native like vowels, the participants need to have had extensive L2 experience which has been reported to have a more important role on the production than on the perception of L2 new categories (Bohn and Flege 1990).

As for CG contrasts, they matched PAM predictions with their very good identification and discrimination patterns, except English (KIT, DRESS) vowels. The CG contrasts were more difficult for SA participants in perception, particularly the English (KIT, DRESS) contrast. This contrast had poor identification and discrimination; nevertheless, the two vowels were produced acoustically distinctly from each other. However, their productions were closer to SA/L1 vowels /i/ and /e/, these SA vowels seem to merge into a single category that is closer to the NE KIT vowel. A full analysis of the SA vowel system is needed to confirm this finding, which is the aim of the following chapter. The SA participants had little difficulty identifying and discriminating the English (LOT, STRUT) and (GOAT, THOUGHT) vowels, and the members of the two contrasts were produced as acoustically distinct from each other.

Finally, the English UC contrast (NEAR, FLEECE) had excellent identification for the categorised FLEECE vowel and good for the uncategorised NEAR vowel. Additionally, both had very good discrimination which matches PAM predictions. In production, L2 productions of the English FLEECE and NEAR vowels were acoustically distinct from each other, but their productions did not match NE nor L1 vowels.

In general, SA participants seem to have no great difficulty in perceiving most of the English vowels under investigation. For the vowel contrasts in Table 3-11 above, the participants managed to identify and discriminate most of the vowels either by referring to a similar/close L1 vowel contrast, or by referring to perceptual phonetic distances between a particular part of vowels, and this was particularly the case with diphthongs. For most of the vowel contrasts in Table 3-11, the members of these contrasts were produced as acoustically distinct (DP), except the (FACE, SQUARE) vowel contrast. However, the production patterns of most English vowels under investigation were acoustically different from native English productions, except for the TRAP vowel. Perhaps, it is more important to be understood in a classroom context than to be native like. For this reason, learners would care more to have distinct categories than native like productions. Thus, in a more fine grained production analysis, the SA productions of the above vowels could be described as "DP-not native like".

3.5 Pilot Study Limitations

The pilot study had some limitations; some of which were avoided in the main study. First, the scope of the pilot study was to cover the full vowel system of SSBE and SA, which was a great amount of work. Also, having three experiments with the full set of vowels made each experiment lengthy for participants. For this reason, only two participants from each language were used. However, exploring the perception and production of all vowels of the two languages was extremely useful to narrow down the main study and to formulate testable hypotheses and predictions based on realistic data.

For the identification experiment, it was difficult and in some cases impossible to find minimal pairs to cover all the possible criteria used to examine the identification of all English vowels. Also, normal speech was used as stimuli for this task, so it was impossible to control the different vowel features. For example, if vowel duration was examined, it was impossible to control vowel place and height differences. However, since the scope of the main study is limited to the six vowel contrasts in Table 3-11 above, it was easier to find minimal pairs and to control for the target vowel features.

Additionally, a spelling effect may have affected the results of the identification task. For example, some of the SA words were produced in Standard Arabic; and this alters the target vowels. Additionally, in the identification experiments the English words were given in written form; so, SA participants might have relied on spelling to identify words, or sometimes over generalising spelling-to-sound might have affected their choices. In order not to rely entirely on spelling, pictorial Arabic and English choices were given simultaneously with the spelling.

In the production experiment, a PPT slide-show on a computer screen was used to present the stimulus material for the pilot study; however, it was difficult to avoid the computer fan noise. Thus, flash cards were used instead of a PPT slide-show for the main study to get clear sound files without any noise from the computer.

3.6 Conclusions

For the main study in Chapters 5-7, the English contrasts set out in Table 3-11 were re-examined with a greater number of SA and NE participants. The perceptual patterns of these vowels were examined in a PAT similar to the pilot study, with the addition of the SA allophonic form [aːˤ] which is predicted to map to English BATH vowel, and thus, the (TRAP, BATH) vowel contrast were examined as a Two Category (TC) assimilation type rather than as SC. Also, the identification task was redesigned to examine the identification and discrimination of the above contrasts in particular. The productions of the above vowels were examined with the addition of the SA allophone [aːˤ] to the list of SA vowels to be compared to the productions of the English BATH vowel.

In the next chapter, the SA vowel system was analysed acoustically, in order to determine the status of its phonemic and phonetic categories. The implications of the results of this analysis on L2 perception and production were taken into consideration in the analysis of the results of the main study in Chapters 5-7. The analysis of the results of the pilot study revealed that it is important to have an acoustic analysis of the SA vowel system including mid vowels and the schwa. In the literature, the SA vowel system has been auditorily described (Cowell 1964), and the high and low vowels have been acoustically examined (Allatif 2008, Allatif and Abry 2004, Munro 1993). No previous research has been dedicated to fully explore the vowel system of SA. The results of the present study showed that long mid vowels play an important role in the perception and production of some English vowels, particularly diphthongs. Similarly, the acoustic analysis showed that SA short mid vowels overlap, and this finding had a significant effect on the perception and production of English short vowels, particularly the (KIT, DRESS) vowel contrast. Thus, it is critical to analyse SA short vowels, in general, and mid short vowels, in particular.

All in all, the pilot study proved to be extremely informative in terms of examining a larger data set on a smaller number of participants, which provided some answers to the research questions of each of the three experiments. Thus, the same questions were carried forward to the main study for further examination with more controlled data and more participants from each language. The analysis of the data in the main study was based on concrete set of assimilation, identification, discrimination, and production categorisations. For example, although it was not specified in PAM how to classify a particular contrast (as SC or TC ... etc), the main study of the present thesis developed a scale to set up PAM categories. Similarly, a scale was developed to set up the identification, the discrimination, and the production results.

4 Chapter 4: Acoustic Analysis of the Syrian Arabic Vowel System

This chapter presents an acoustic analysis which has been conducted to examine the phonetic properties of the full vowel system of SA. The main motivation for the present study was the need to have a full acoustic analysis of the vowel system for the first language (L1) of the participants. No previous research has been dedicated to examine the vowel system of SA, except for the auditory description of Cowell (1964) which was based on one male Damascene speaker only. On the other hand, many studies (Deterding 1997, Hawkins and Midgley 2005, Hughes et al. 2005, Wells 1986) have been dedicated to examine the SSBE vowel system, which is the second language (L2) of the participants. Thus, the results of these studies can be used as a reference to be compared to SA participants' L2 productions. Nevertheless, it is important to run an acoustic analysis of SA vowels, in order to have a quantitative study of L1 (SA) comparable to those of SSBE. Additionally, this study is needed to examine the effect of L1 (SA) on the participants' L2 perceptual and production patterns.

The present chapter is organised as follows: first, section 4.1 presents the stimulus material, the participants, and the methods used to collect the SA data. Section 4.2 presents the details of the acoustic analysis procedures followed for this study. Section 4.3 presents the overall results of the acoustic analysis of SA vowels, which included a statistical examination of the general effects on the whole data set. In Section 4.4 pairwise comparisons of the SA short-long vowel pairs are presented, in order to determine the quantitative and qualitative relationships between these vowel pairs. After that, section 4.5 presents analysis of the acoustic overlap which has been found among SA short vowels. The main conclusion of this analysis is that SA has a three short vowel system consisting of /i/, /a/, /u/. The short mid vowels /e/, /o/ are analysed as phonetic variants/allophones of their high counterparts, and the schwa as a phonetic variant/allophone which might surface in place of any other SA short vowel. A summary of the results of this study is presented in section 4.6. Finally, a discussion of the main results and findings is given in section 4.7, followed by considering the implications of the results of this chapter on the participants' perceptual and production patterns of English vowels as a second language, in section 4.8.

4.1 Methods

The focus of this analysis is, as throughout this thesis, the variety spoken in Damascus, the capital of Syria. Ten male and ten female SA speakers participated in this study, but due to technical problems with the recordings, the productions of the ten male and five of the ten female participants were analysed. All participants were from Damascus. On average, the male participants were 23.5 years old and the female participants were 30.6 years old. The participants were asked to complete a questionnaire about their language background; more details about this questionnaire will be presented in the following chapter.

The participants were recorded at the Asia Institute for Languages in Damascus. They were recorded in a quiet computer room using a Marantz Professional Portable Solid State Recorder, PMD660 model. The microphone was Shure Professional unidirectional, which was a head-worn dynamic microphone. Audio files were recorded at 16 bit and 44.1 Khz sampling rate as .wav files on a compact flash TM memory card.

Cowell (1964) claims that the vowel system of SA consists of five long vowels (/iː/, /eː/, /aː/, /oː/, /uː/), five short vowels (/i/, /e/, /a/, /o/, /u/), and a schwa /ə/. Table 4-1 below presents an attempt to list minimal pairs for Cowell's suggested vowel categories. Some of these words were extracted from Cowell (1964). Table 4-1 is an attempt because Cowell's descriptions of SA vowels did not include a straightforward presentation of minimal/near minimal pairs of the suggested vowel categories. The table shows that there are minimal pairs to represent the five short-long vowel pairs, which supports the view that these contrasts are phonemically established in SA. Additionally, front and back long mid-high ((/iː/-/eː/), (/uː/-/oː/)) and long mid-low vowel pairs ((/eː/-/aː/), (/oː/-/aː/)) have minimal pairs to represent these categories.

As for the short vowels, there are minimal pairs in SA for the high (/i/, /u/) and low /a/ short vowels which supports their phonemic status. To my knowledge, there are no minimal pairs to represent any contrast involving mid short vowels (/e/, /o/) or /ə/. Generally, Arabic speakers recognise the short high (/i/, /u/) and low /a/ vowels as phonologically functioning, which is a direct effect of Standard Arabic. The difficulty to find minimal pairs including the short mid vowel categories, suggested by Cowell, can be assumed to be an effect of the Standard Arabic short vowel phonological system. As a result, SA speakers (including me) identify mid short vowels as their high counterparts of Standard Arabic /i/ or /u/, respectively. This assumption also applies to the mid short vowels forming short-long mid vowel contrasts (/e/-/eː/) and (/o/-/oː/), these are given in the grey cells in Table 4-1 below. From a native listener's point of view, the short vowels in these contrasts may alternate between mid and high short vowel qualities.

Table 4-1: SA minimal pairs, some of which were chosen from Cowell (1964), representing vowel phonemic distinctions

vowel	Minimal pairs			
pair	SA word	Gloss	SA word	Gloss
(/i/-/iː/)	[sidd]	'close'	[siːd]	'master'
(/a/-/aː/)	[malek]	'king'	[maːlek]	'owner'
(/u/-/uː/)	[kura]	'ball'	[kuːra]	'name of a village'
(/e/-/eː/)	[baːret]/ [baːrit]	'she competed'	[baːreːt]	'I competed'
(/o/-/oː/)	[tˤabbon]/ [tˤabbun]	'he dropped them down'	[tˤabboːn]	'the front part of the car'
(/iː/-/eː/)	[biːdˤ]	'white-blural'	[beːdˤ]	'eggs'
(/iː/-/aː/)	[fiːʔ]	'wake up!'	[faːʔ]	'he woke up'
(/iː/-/uː/)	[diːb]	'wolf'	[duːb]	'melt'
(/iː/-/oː/)	[fiːʔ]	'wake up!'	[foːʔ]	'up'
(/uː/-/oː/)	[duːr]	'turn'	[doːr]	'queue'
(/uː/-/eː/)	[sˤuːf]	'wool'	[sˤeːf]	'summer'
(/uː/-/aː/)	[suːd]	'black-plural'	[saːd]	'he lead/ spread'
(/eː/-/aː/)	[tˤeːf]	'shadow'	[tˤaːf]	'he goes around'
(/eː/-/oː/)	[deːr]	'abbey'	[doːr]	'queue'
(/oː/-/aː/)	[shoːf]	'vision'	[shaːf]	'he saw'
(/i/-/a/)	[ridd]	'you respond!'	[radd]	'he responded'
(/i/-/u/)	[naːdi]	'club'	[naːdu]	'they called'
(/u/-/a/)	[naːdu]	'they called'	[naːda]	'he called'

The stimulus material for this study is the same as those used for the SA data in the pilot study (chapter 3). The pharyngealised SA phonetic vowel category [aːˤ] was added to the analysis of the SA vowel system, in order to complement the overall goal of the present study, which is examining the effect of the SA vowel system on SA participants' L2 perception and production of English vowels. The analysis of SA vowel system included Cowell's suggested SA phonological vowel categories and the phonetic [aːˤ] category which was suggested, in the pilot study, to have a role in the perception and production of the English BATH vowel. All SA vowels were produced in the /hVd/ context, which was inserted in the phrase /ktoːb ___ marteːn/ "Write ___ twice". The SA vowels were extracted and analysed following the same procedures of the pilot study. Together, the SA stimuli included three repetitions × 12 vowels × 15 speakers = 540 items.

4.2 Data Analysis

The acoustic analysis of the SA vowels included extracting vowel duration and midpoint F1 and F2 frequency measurements. Similar to the pilot study, the measurements were extracted automatically using a Praat script. However, the landmarks for the beginning and the ending of the vowels were marked by hand using textgrid files. Additionally, the data was carefully checked for outliers, particularly for the frequency measurements. The measurements of the detected outliers were checked, analysed, and corrected by hand in Praat.

In order to normalise the effect of speech rate on vowel duration, the duration of the vowel was normalised over the duration of the phrase "say /hVd/ again" for each vowel. The following formula (1) was used to calculate normalised vowel duration:

(1) Normalised vowel duration = ('vowel dur'/'phrase dur') * 100

As explained in Chapter 2, the anatomical differences of the individual speakers and those of male and female speakers would alter the frequency measurements of the same vowels. Thus, the frequency measurements of F1 and F2 were normalised using the LOBANOV procedure which was found to succeed in preserving the interesting phonemic and sociolinguistic variation and minimising the anatomical variation (Adank, et al. 2004, Fabricius, et al. 2009). In a recent study, Flynn (2011) evaluated the efficiency of 20 normalisation procedures at neutralising inter-speaker anatomical and physiological differences. He also found that LOBANOV performed well compared to the majority of those procedures. Thus, the LOBANOV procedure which is implemented in an online software package NORM (Thomas and Kendall 2007), was used to normalise the F1 and F2 values of the present study.

4.3 Overall Results

Figure 4-1 below presents mean vowel durations for the SA vowels as produced by male and female speakers. As can be seen in the figure, there are no apparent differences between male and female productions in terms of vowel durations, which is similar to what has been found for Lebanese Arabic male and female absolute and proportional durational differences (Khattab and Al-Tamimi 2008). The figure also shows that the mean duration of the long vowels is almost double that of the short vowels, which confirms the predicted durational difference between short and long vowel contrasts in SA. Additionally, the duration of the schwa which does not have a long counterpart is produced shorter than the other short vowels.

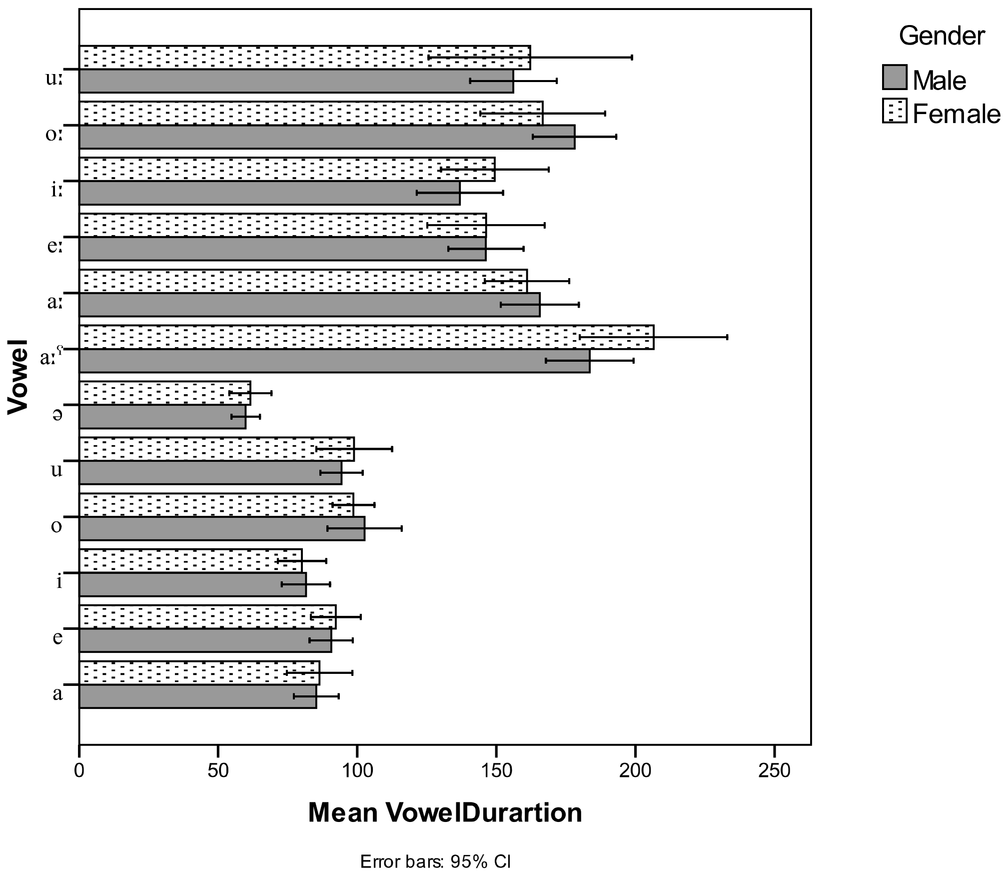

Figure 4-1: Male and female mean vowel duration values of the SA long vowels (/iː/, /eː/, /aː/, /oː/, /uː/), short vowels (/i/, /e/, /a/, /o/, /u/), the schwa /ə/, and the allophone [aːˤ] based on the auditory description of Cowell (1964)

Table 4-2 below presents the ratio of short to long vowel pairs for male and female productions. The table shows that the overall ratio of V to VV is 1 to 1.71, which is less than what has been reported for Lebanese Arabic: 1 to 2.02 (Khattab and Al-Tamimi 2008). The duration of SA long vowels is less than double that of their short counterparts, which is less than what is expected. Additionally, the ratio of V to VV is the same for the productions of both males and females.

Table 4-2: Ratio of male and female productions of SA short-long vowel pairs, the ratios were calculated based on actual vowel durations in ms

Short-long vowel contrasts	Males' (V-VV) ratio	Females' (V-VV) ratio	Overall (V-VV) ration
/i/-/iː/	1:1.67	1:1.87	1:1.77
/e/-/eː/	1:1.6	1:1.58	1:1.59
/a/-/aː/	1:1.95	1:1.87	1:1.91
/o/-/oː/	1:1.72	1:1.68	1:1.7
/u/-/uː/	1:1.65	1:1.63	1:1.64
Average	**1:1.71**	**1:1.72**	**1:1.71**

Vowel duration differences between SA vowels were examined statistically using a linear mixed model, in which 'vowel type' (SA 11 vowels) and 'sex' (M/F) were included as fixed effects and 'participant' (15 SA participants) was included as a random effect. As predicted, the results showed that there was a significant effect for 'vowel type' on vowel duration differences: $F(11,509) = 71$, $p < .000$. However, there was no significant effect for 'sex' $F(1,509) = .666$, $p = .415$, nor for 'vowel type' and 'sex' interaction: $F(11,509) = .730$, $p = .710$.

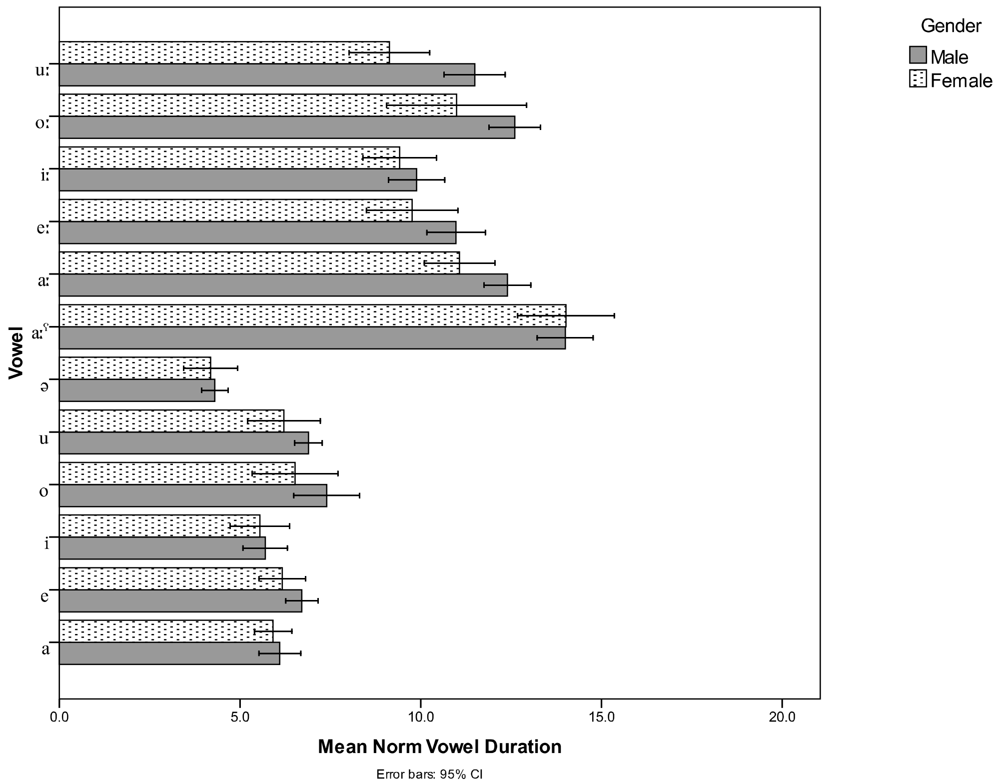

Figure 4-2: Male and female mean normalised vowel duration values of SA long vowels (/iː/, /eː/, /aː/, /oː/, /uː/), short vowels (/i/, /e/, /a/, /o/, /u/), the schwa /ə/, and the allophone [aːˁ] based on the auditory description of Cowell (1964)

The effect of speech rate on SA vowel duration was normalised over the phrase duration, as explained in the previous section 4.2. The normalised vowel durations are presented in Figure 4-2 above, which shows essentially the same patterns as Figure 4-1. The ratios of short to long vowel pairs were calculated again in Table 4-3 below based on the normalised values of vowel durations. The table shows that the overall V to VV ratio (1:1.63) is less than the ratio which was calculated without normalising vowel duration (1:1.71). Thus, the duration of SA long vowels can be argued to be approximately one and a half time longer than SA short vowels.

This later finding matches the ratio of V to VV which was calculated for Lebanese Arabic V1CV2 and VV1CV2 contexts in spontaneous speech (1:1.58) (Khattab 2007). However, the two ratios of SA and Lebanese Arabic were less than the ratio of V to VV calculated for word list in Lebanese Arabic productions in Khattab (2007) (1:2.08) and Khattab et al. (2008) (1:2.01). Khattab (2007), which was dedicated to examining the phonetic correlates of gemination in Lebanese Arabic, reported that the ratio of geminated to singleton consonants in spontaneous speech (1:1.82) was also less than the ratio in the word list productions (1:2.29). Additionally, the ratio of V to VV reported in the present study (1:1.63), in Khattab (2007) (1:2.08) and Khattab et al. (2008) (1:2.02) is less than the ratio reported for C to CC (1:1.97, calculated across the later two studies).

Table 4-3: Ratio of male and female productions of SA short-long vowel pairs, the ratios were calculated based on the normalised vowel durations values

Short-long vowel contrasts	Males' (V-VV) ratio	Females' (V-VV) ratio	Overall (V-VV) ration
/i/-/iː/	1:1.6	1:1.5	1:1.6
/e/-/eː/	1:1.5	1:1.6	1:1.5
/a/-/aː/	1:2	1:1.83	1:1.91
/o/-/oː/	1:1.85	1:1.5	1:1.67
/u/-/uː/	1:1.5	1:1.5	1:1.5
Average	**1:1.69**	**1:1.58**	**1:1.63**

The differences between SA vowels in terms of the 'normalised vowel duration' was also examined using a linear mixed model with 'vowel type' and 'sex' as fixed effects and 'participant' as a random effect. The results again showed that there was a significant effect for 'vowel type' on 'normalised vowel duration' differences: F (11,509) = 107, p < .000. However, with the normalised data 'sex' had a significant effect: F (1,509) = 21.4, p < .000. Generally, normalised vowel durations of males' productions were longer than females'. However, there was no significant effect for 'vowel type' by 'sex' interaction: F (11,509) = 1.397, p = .170. The results of the normalised vowel duration will be presented and statistically examined for all participants from both sexes.

As for quality, Figure 4-3 and Figure 4-4 below present the normalised F1 and F2 frequency values for SA female and male speakers, respectively. The frequency measurements were normalised following the LOBANOV normalisation procedure explained in the previous section. Apart from the allophonic [aːˤ] vowel category, the figures show that females and males have the same general triangular shape reported for Standard Arabic (Newman and Verhoeven 2002), with the addition of more centralised mid vowel categories. Figure 4-3 shows that females' productions exhibit more within category variation than males' productions in Figure 4-4. Additionally, the central area of the vowel space appears vacant for males' productions whereas some of the females' productions of the schwa and /i/ vowels were more centralised.

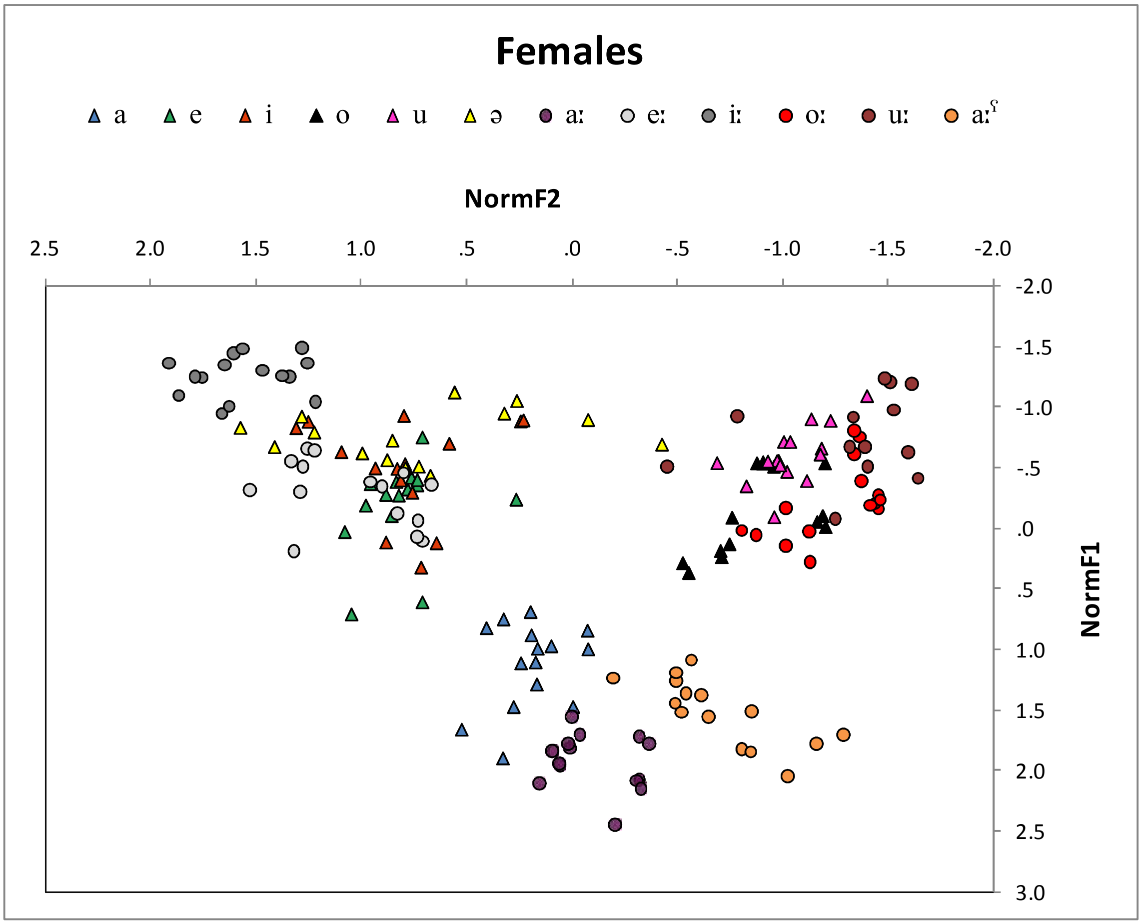

Figure 4-3: Females' LOBANOV Z-scores of the normalised F1 and F2 values of SA vowels based on the auditory description of Cowell (1964)

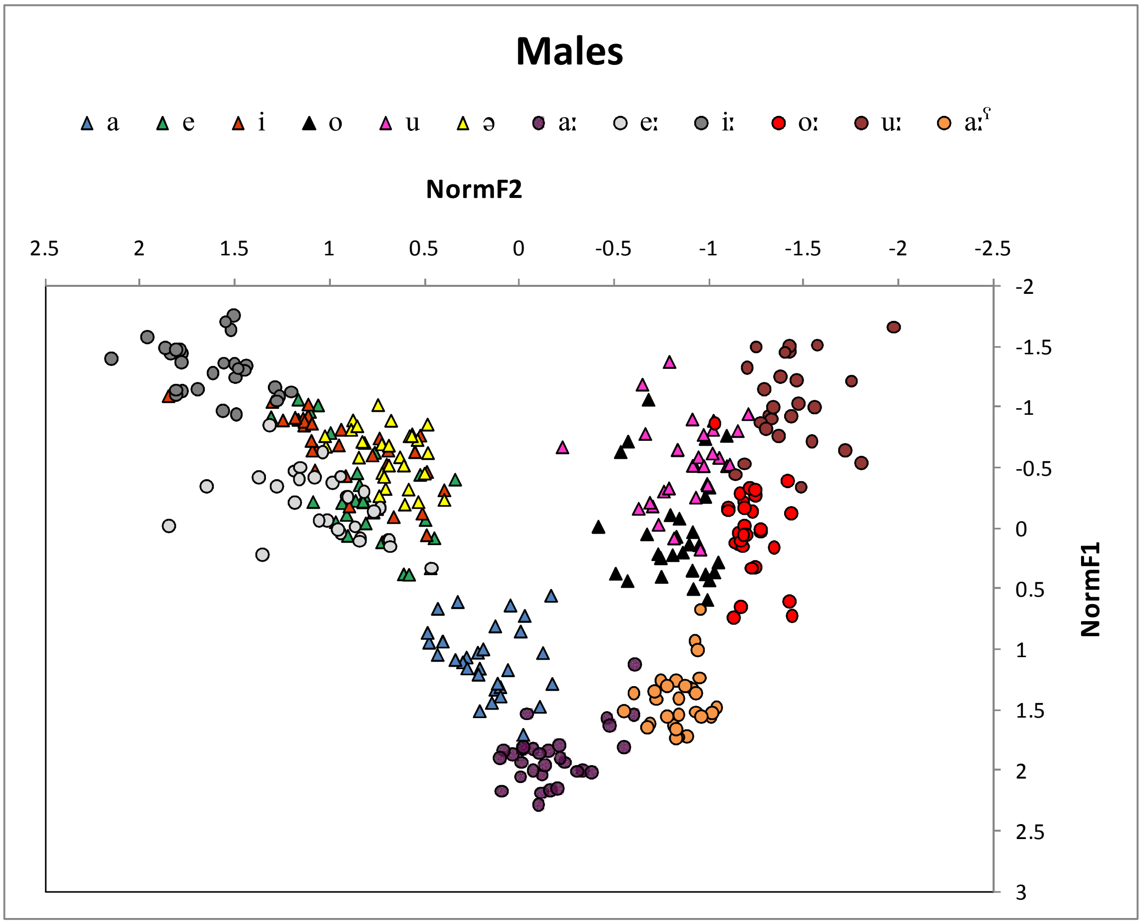

Figure 4-4: Males' LOBANOV Z-scores of the normalised F1 and F2 values of SA vowels based on the auditory description of Cowell (1964)

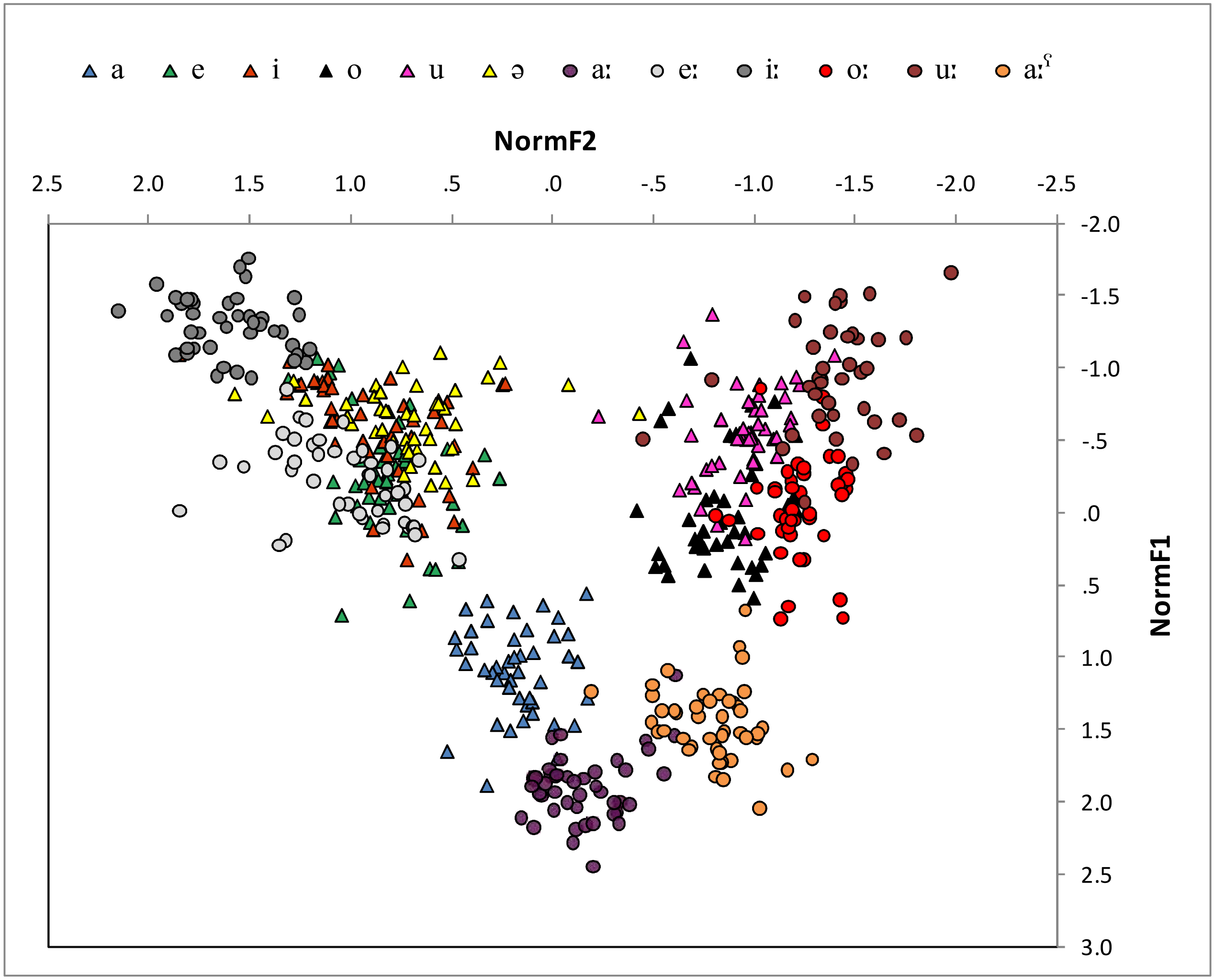

Figure 4-5: LOBANOV Z-scores of the normalised F1 and F2 values of all SA vowels for all SA participants

Figure 4-5 above presents a scatter plot of the normalised F1 and F2 values for all the SA vowels across all the participants. Apart from the few schwa and /i/ tokens of the female speakers, the chart shows that the central area of the vowel space is almost empty and not occupied by any vowel, which might lead us to question the status of the central vowel, i.e. the schwa, in SA.

As can be seen in the chart, the short vowels appear more centralised than their long counterparts. A closer examination to short-long vowel contrasts shows that there is a clear separation in the vowel space between SA (/iː/, /i/), (/aː/, /a/), and (/uː/, /u/). However, there is no clear separation in the vowel space between SA short-long mid vowels (/eː/, /e/) and (/oː/, /o/). Additionally, there is a great overlap in the mid-high area in the vowel space, particularly between short vowels (/i/, /e/, /ə/) and (/u/, /o/). The overlap between these categories in the vowel space as well as in duration suggests that the difference between these categories might be purely phonetic rather than phonological. Particularly, that it was difficult to compose minimal pairs involving the short mid vowels /e/, /o/, and the schwa in Section 4.1 above.

The pharyngealised allophonic [aːˤ] category appears in the vowel space more retracted and slightly higher than its plain cognate /aː/, thus, corroborates the results in the literature of pharyngealisation (Almbark 2008, Khattab, et al. 2006). Interestingly, the phonetic difference between the plain/pharyngealised (/aː/, [aːˤ]) contrast is more prominent than the phonetic differences between the overlapping short vowels, which are claimed to be phonological categories in SA (Cowell 1964). Nevertheless, the difference between the (/aː/, [aːˤ]) categories is agreed to be phonetic and does not have any phonological role in Arabic. This agreed upon fact adds to the argument mentioned earlier that the difference between the overlapping short vowels is purely phonetic and does not have any phonological function in the distinction of these categories.

Two Additional linear mixed model tests were conducted to examine the differences between SA vowels in terms of the ‘F1’ and ‘F2’. The results showed that there was a significant effect for ‘vowel type’ on ‘F1’ differences: $F\ (11,509) = 395.9$, $p < .000$, and ‘F2’ differences: $F\ (11,509) = 708.5$, $p < .000$. On the other hand, ‘sex’ had no significant effect on ‘F1’ differences: $F\ (1,509) = .025$, $p = .873$, nor on ‘F2’ differences: $F\ (1,509) = .001$, $p = .972$. The interaction of ‘vowel type’ and ‘sex’ had no significant effect on ‘F2’: $F\ (11,509) = 1.433$, $p = .154$, however, the interaction had a significant effect on ‘F1’ differences: $F(11,509) = 2.162$, $p = .015$. Further examination of the data showed that males had significantly lower (/o/, /u/, /ə/, /eː/, and /oː/).

In section 4.4, the results of Bonferroni post hoc tests will be presented for further analysis of the paired comparisons of SA short-long vowel pairs, which will be compared in terms of quantity and quality. After that, the overlap of the short vowels in the mid-high areas of the vowel space will be examined thoroughly in section 4.5.

4.4 SA Short-long vowel pairs

4.4.1 SA high front vowel contrast: (/i/, /iː/)

In this section, the quantitative and qualitative results of SA high front (/i/, /iː/) vowels are presented and compared. Figure 4-6 below presents mean normalised vowel duration of SA (/i/, /iː/) vowels, which shows that SA /i/ is shorter in duration than SA /iː/ (a ratio of 1 to 1.6). A Bonferroni post hoc test confirmed that SA /i/ vowel is significantly shorter ($p < .000$) than its long counterpart.

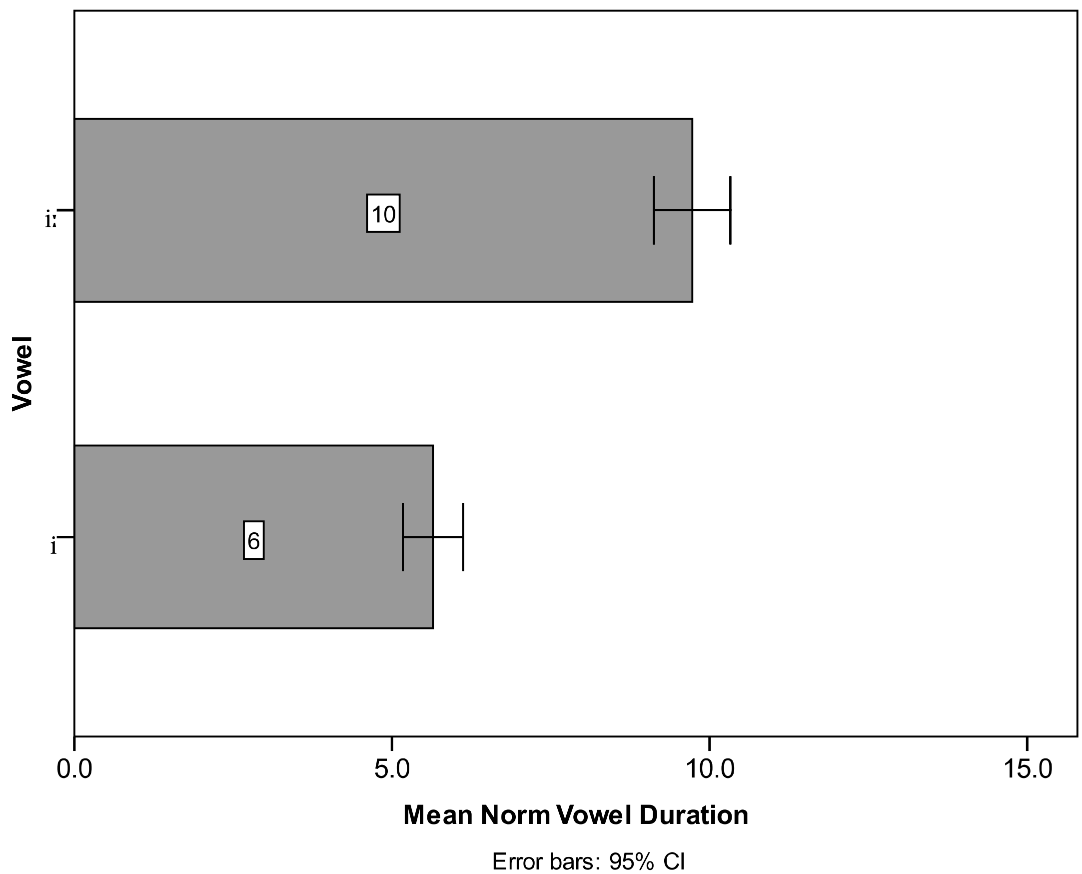

Figure 4-6: Mean normalised vowel duration of SA short-long vowel contrast (/i/, /iː/) as produced by SA participants

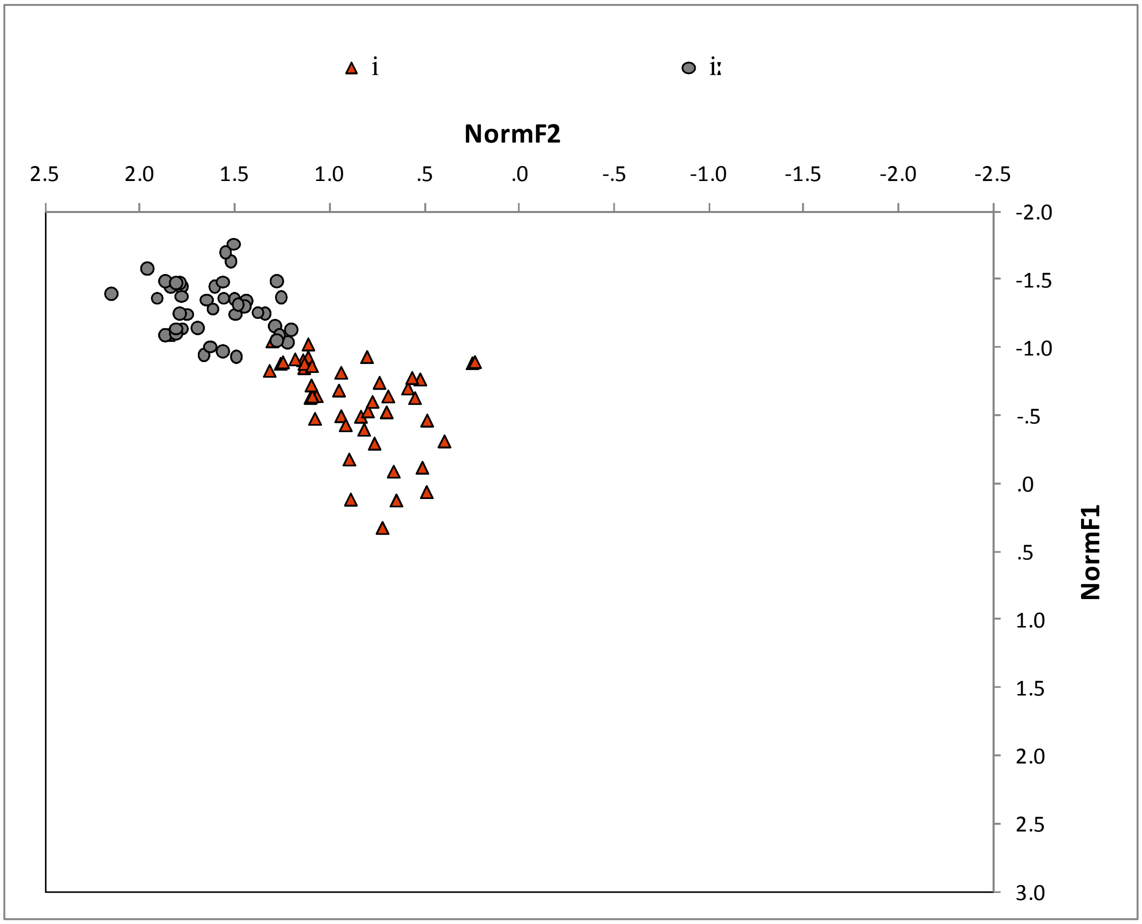

Figure 4-7: scatter plot of normalised F1 and F2 values of SA short-long vowel contrast (/i/, /iː/), using LOBANOV normalisation procedure

Figure 4-7 above presents a scatter plot of the normalised F1 and F2 values following LOBANOV normalisation procedure. As shown in the figure, the SA short vowel /i/ is predictably lower and more centralised than its long counterpart /iː/. The results of the Bonferroni post hoc tests supported this prediction, in which the SA /i/ vowel was significantly lower (F1: $p < .000$), and more centralised (F2: $p < .000$) than the SA long vowel /iː/. Thus, the members of SA (/i/, /iː/) vowel contrast differ from each other in terms of quantity and quality.

4.4.2 SA high back vowel contrast: (/u/, /uː/)

Figure 4-8 below shows the mean normalised duration of the SA high back vowel contrast (/u/, /uː/), which shows the same pattern of the previous contrast. The results of the Bonferroni post hoc test showed that SA /u/ vowel is significantly shorter than its long counterpart ($p < .000$), with a very close (V to VV) ratio to the previous contrast (1 to 1.5) .

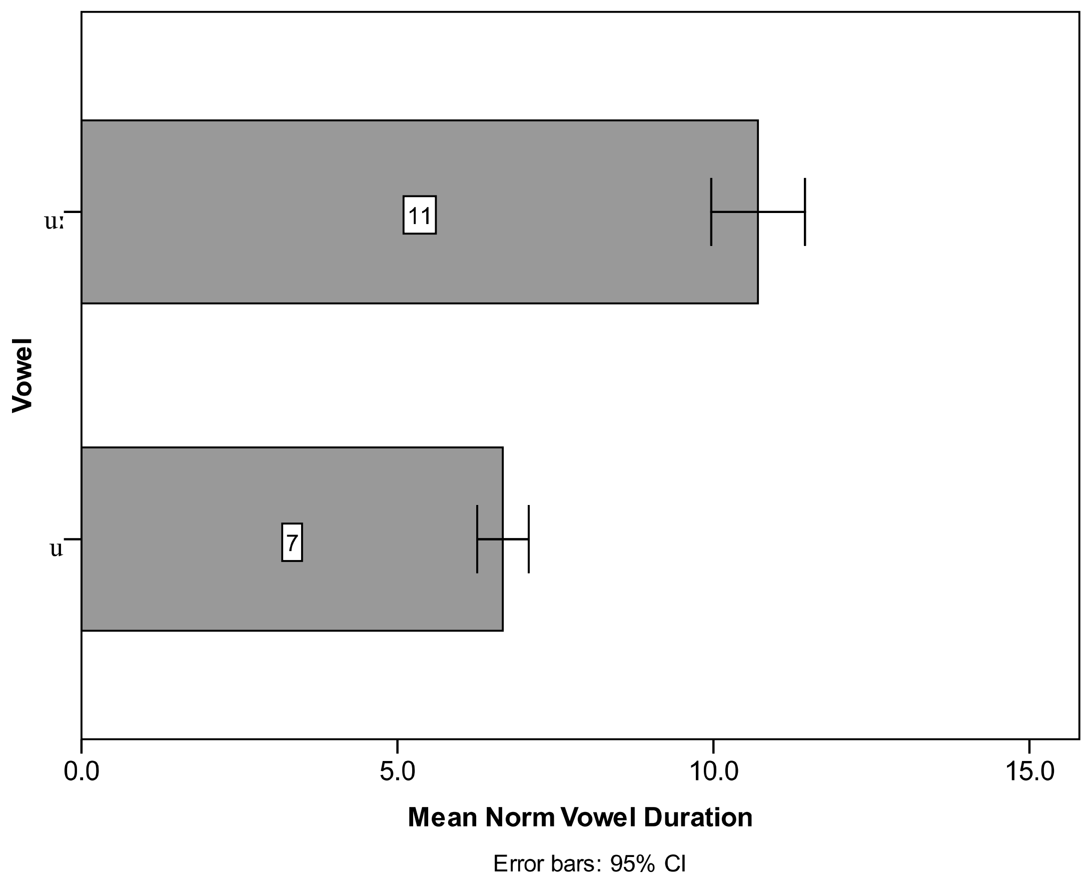

Figure 4-8: Mean normalised vowel duration of SA short-long vowel contrast (/u/, /uː/) as produced by SA participants

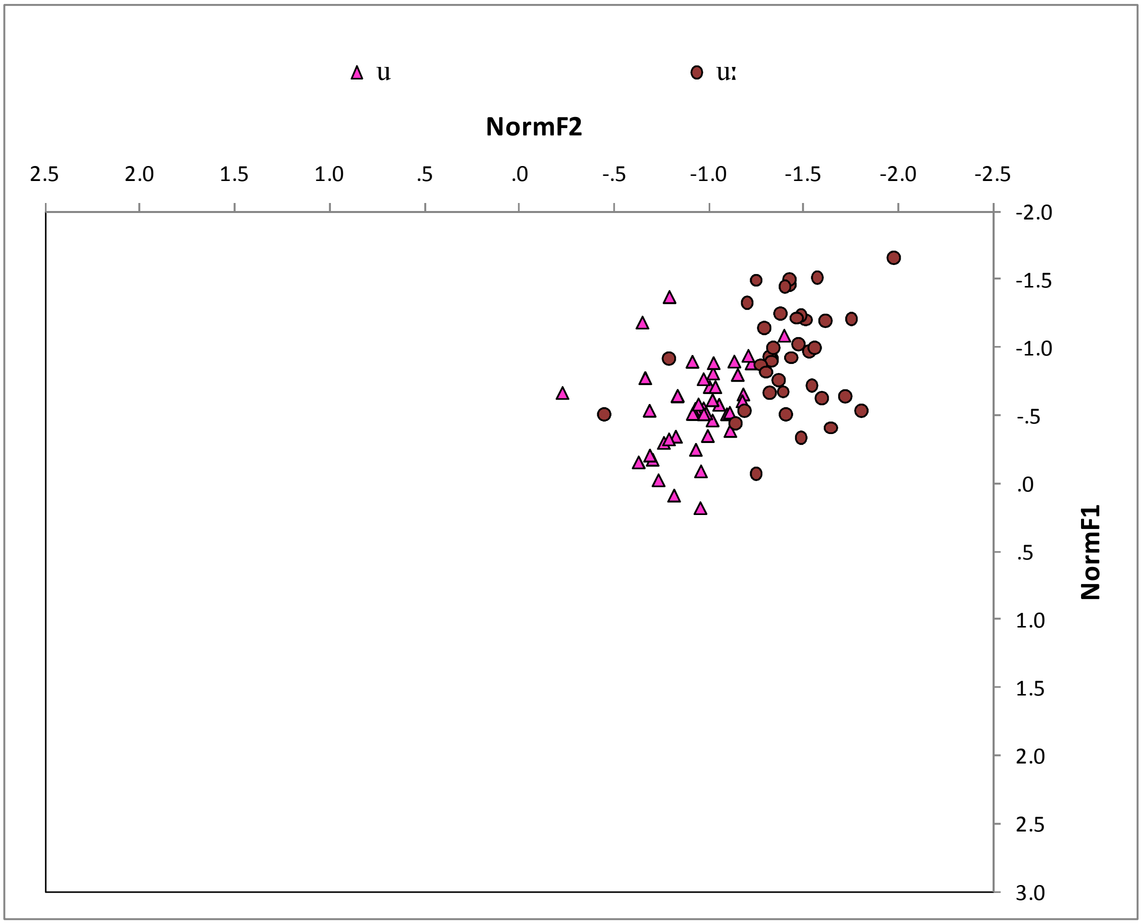

Figure 4-9: Scatter plot of normalised F1 and F2 values of SA short-long vowel contrast (/u/, /uː/), using LOBANOV normalisation procedure

Figure 4-9 above presents the normalised F1 and F2 values of SA (/u/, /uː/) vowels. The figure shows that, similar to the previous high front vowel contrast, the members of the high back vowel contrast (/u/, /uː/) do not overlap in the vowel space. The results of the Bonferroni post hoc tests showed that SA /uː/ vowel was significantly higher (F1: $p = .001$) and more peripheral (F2: $p < .000$) than its short cognate /u/. Thus, the members of the SA high back vowel (/u/, /uː/) contrast differ from each other in terms of quantity and quality, similar to the previous (/i/, /iː/) contrast.

4.4.3 SA low vowel contrast: (/a/, /aː/)

Figure 4-10 below presents the mean normalised duration of SA (/a/, /aː/) low vowels as produced by SA speakers. The results of the Bonferroni post hoc test showed that SA /a/ vowel was significantly shorter ($p < .000$) than its long counterpart /aː/. However, the V to VV ratio of this contrast is slightly higher than the previous contrasts (1 to 1.91), showing that the long vowel is as double as its short counterpart. The ratio for this contrast is close to the ratio calculated for Lebanese Arabic (1 to 2.02) based on Khattab et al. (2008) findings for this contrast. The slightly greater ratio difference of the later study can be due to using word list style without having a carrier phrase. Additionally, vowel durations of the present study were normalised over the duration of the carrier phrase as an attempt to normalise the effect of speech rate.

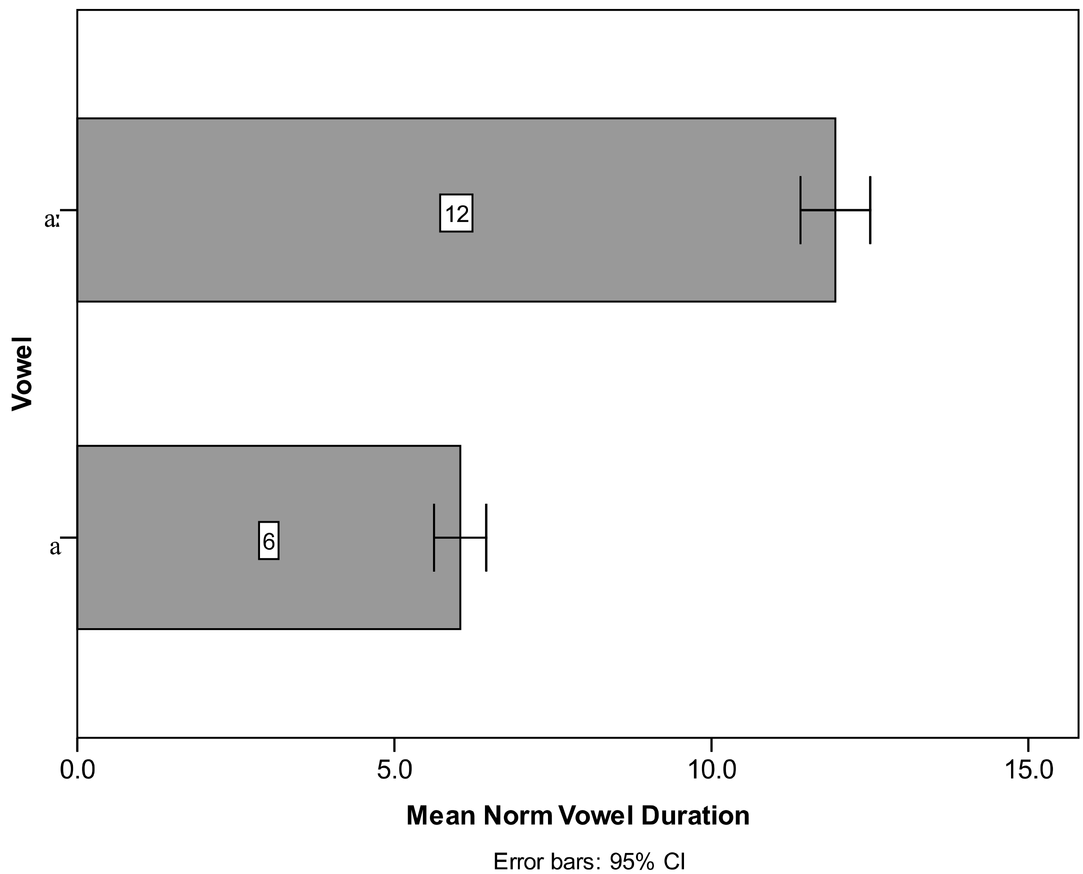

Figure 4-10: Mean normalised vowel duration of SA long-short vowel contrast (/a/, /aː/) as produced by SA speakers

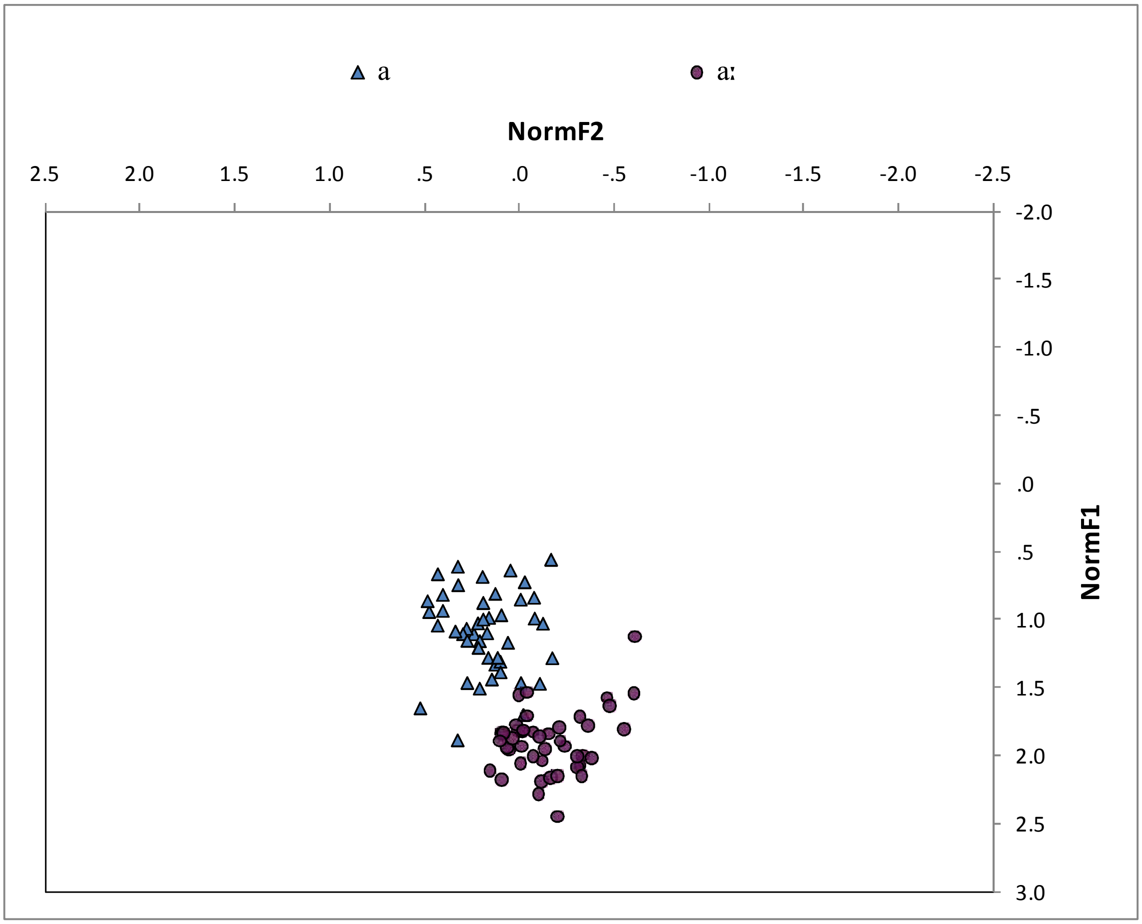

Figure 4-11: Scatter plot of normalised F1 and F2 values of SA long-short vowel contrast (/a/, /aː/), using LOBANOV normalisation procedure

Figure 4-11 above presents the normalised F1 and F2 values for the SA low vowels (/a/, /aː/). As can be seen, the members of the low vowel contrast do not overlap in the F1 dimension of the vowel space; the SA short vowel /a/ appears higher than its long counterpart. However, the two vowels seem to overlap on the F2 dimension for most tokens. The Bonferroni post hoc tests showed that the long vowel /aː/ was significantly lower (F1: $p = .000$) and backer ($p = .000$) than the short /a/ vowel. Thus, the members of the low vowel contrast exhibit a similar pattern to the previous contrasts. In sum, both quantity and quality had significant effects on the phonetic differences between SA high and low short-long vowel contrasts.

4.4.4 SA mid front vowel contrast: (/e/, /eː/)

As for the mid vowels, Figure 4-12 below shows that SA front mid vowel /e/ is shorter in duration than /eː/, showing a V to VV ratio of 1 to 1.5, which is similar to the previous high vowel contrasts. The Bonferroni post hoc test results showed that the long vowel /eː/ was significantly longer ($p < .000$) than its short cognate.

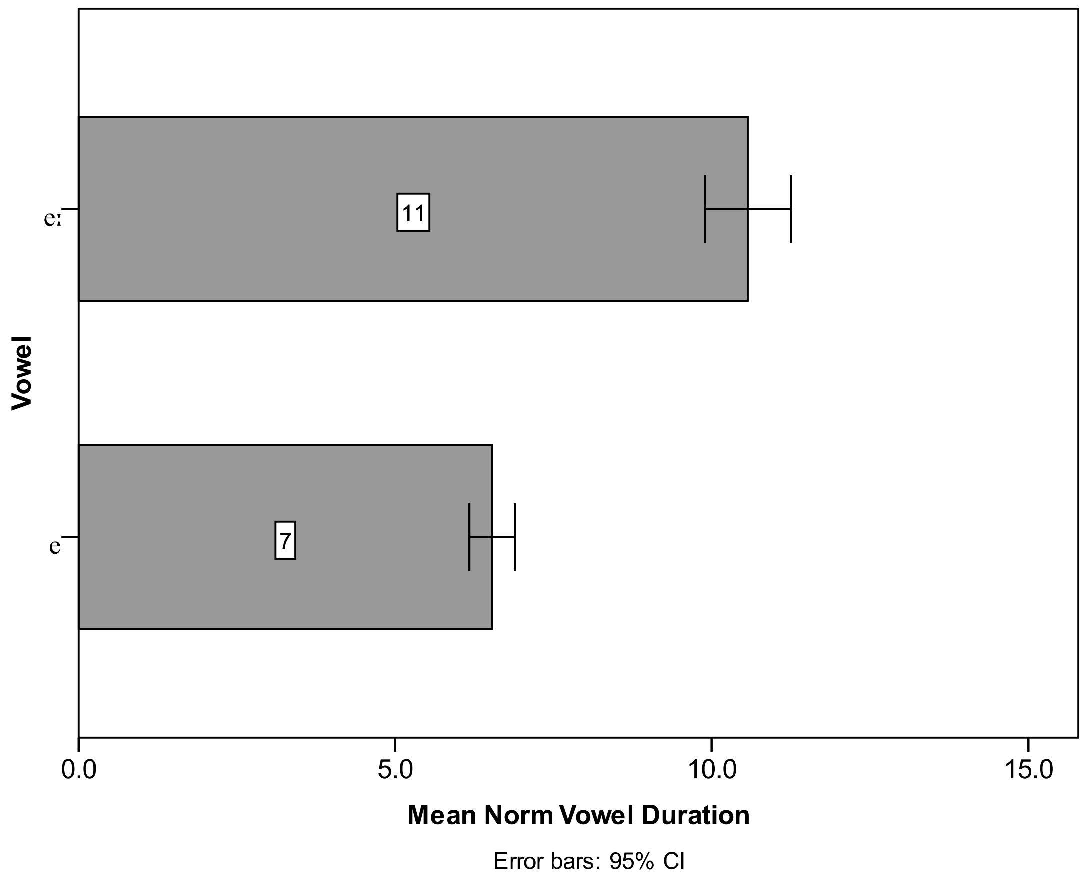

Figure 4-12: Mean normalised vowel duration of SA long-short vowel contrast (/e/, /eː/) as produced by SA participants

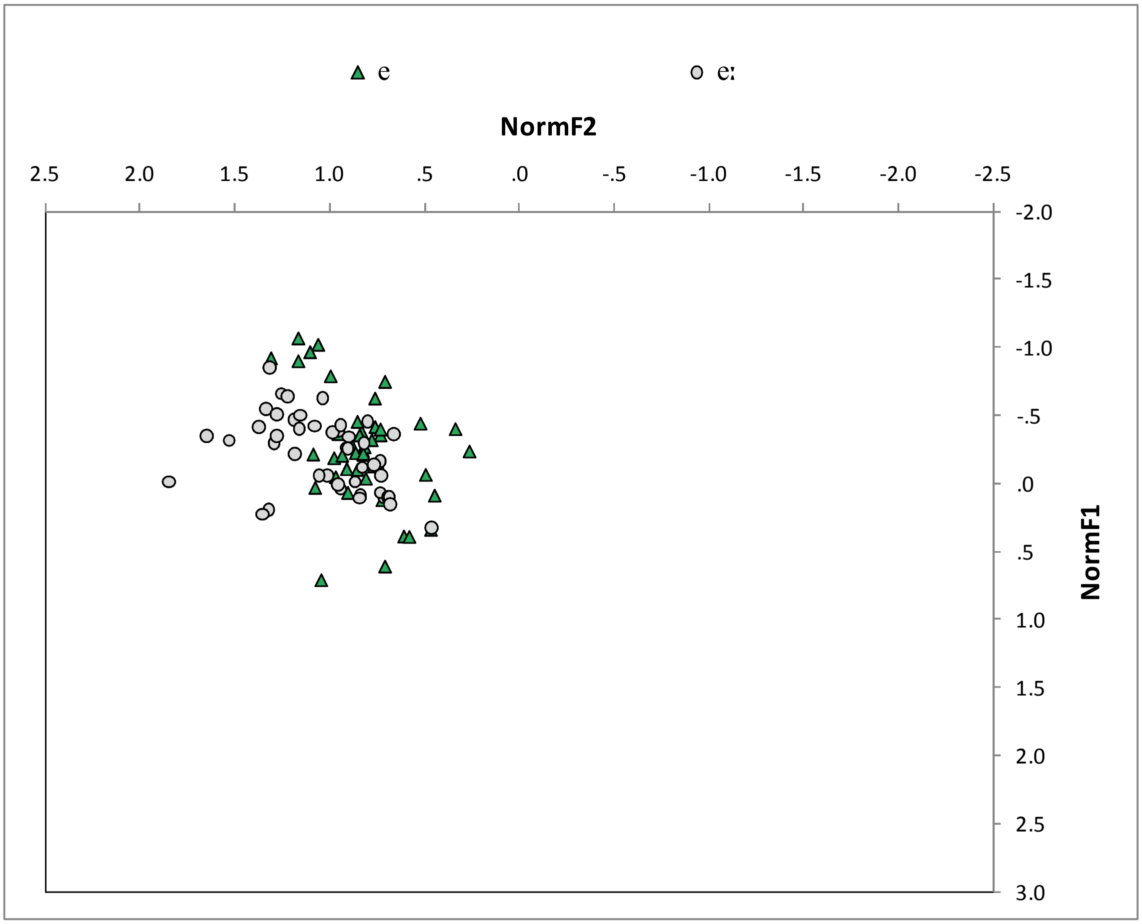

Figure 4-13: Scatter plot of normalised F1 and F2 values of SA long-short vowel contrast (/e/, /eː/), using LOBANOV normalisation procedure

Figure 4-13 above presents a scatter plot of the normalised F1 and F2 values of SA (/e/, /eː/) vowels. Unlike the previous contrasts, there is a great overlap in the vowel space between the long and short SA front mid vowels, which is reflected in the Bonferroni post hoc test results. The long mid vowel /eː/ was significantly fronter (F2: $p = .005$) than its short counterpart /e/, however, both occupied the same area on the F1 dimension (F1: $p = 1$). These results suggest that SA short vowel /e/ is also more centralised than its long cognate, i.e. on the F2 dimension only.

4.4.5 SA mid back vowel contrast: (/o/, /oː/)

The second SA mid vowel contrast (/o/, /oː/) shows a similar durational pattern to the previous contrasts, as shown in Figure 4-14. The results of the Bonferroni post hoc test showed that /o/ had a significantly shorter duration than SA long vowel /oː/ ($p < .000$). Also, the V to VV ratio for this contrast (1 to 1.67) is close to the ratios found for the previous high and mid contrasts.

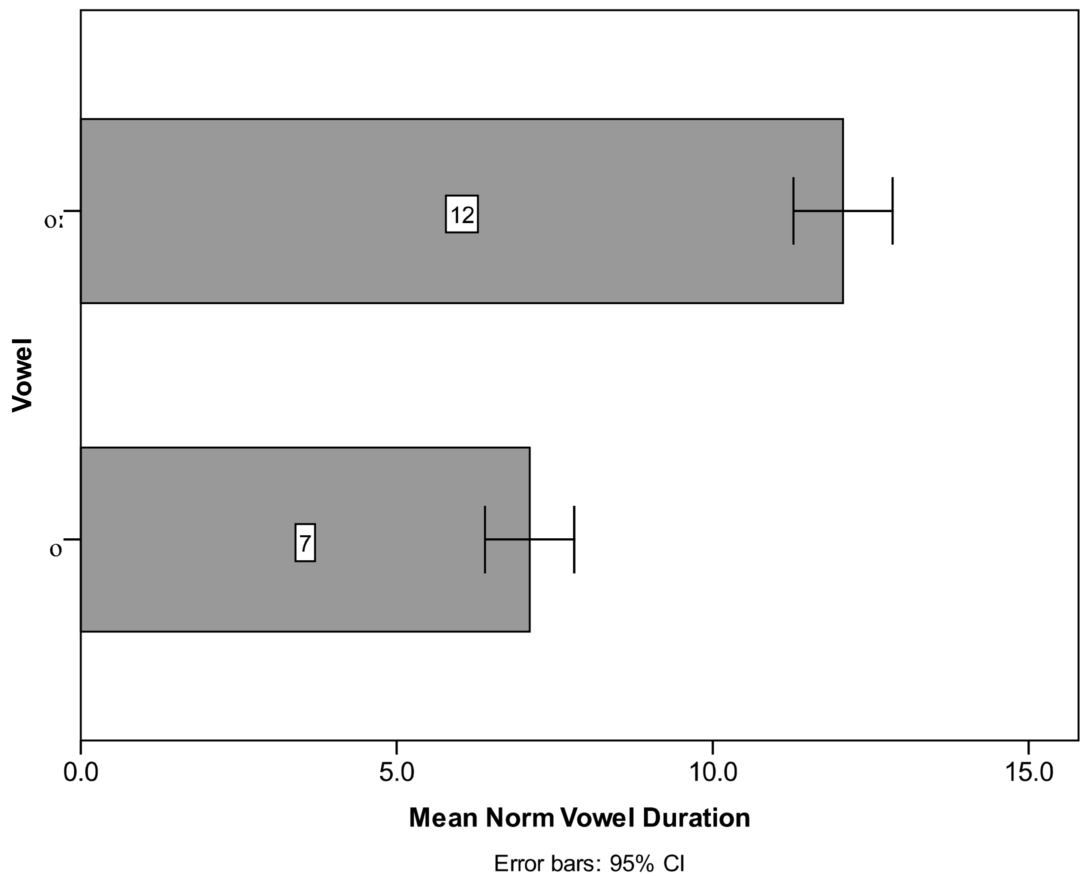

Figure 4-14: Mean normalised vowel duration of SA long-short vowel contrast (/o/, /oː/) as produced by SA participants

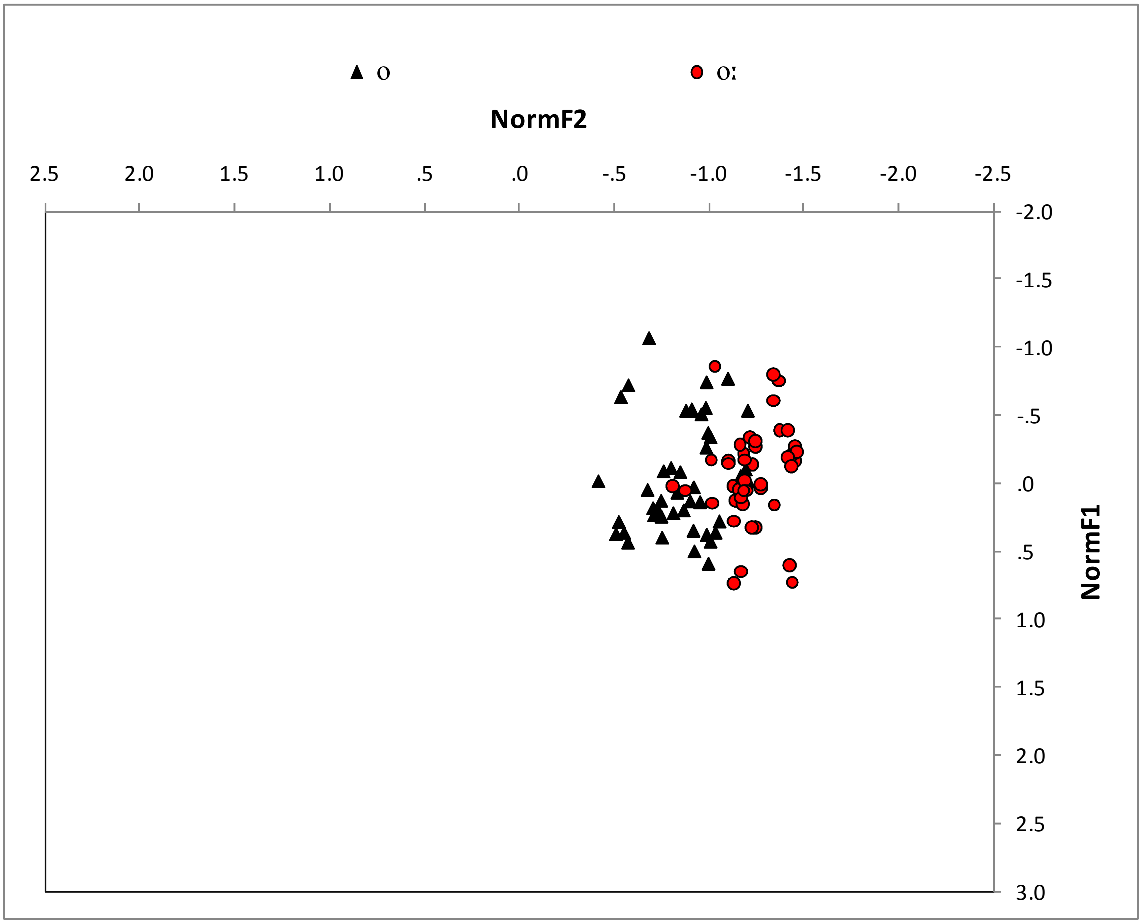

Figure 4-15: Scatter plot of normalised F1 and F2 values of SA long-short vowel contrast (/o/, /oː/), using LOBANOV normalisation procedure

In the vowel space, SA (/o/, /oː/) vowels show a great overlap similar to that of the front mid vowel contrast (/e/, /eː/), as can be seen in Figure 4-15. The Bonferroni post hoc tests showed that the long vowel /oː/ was significantly backer on the F2 dimension ($p < .000$). However, the difference on the F1 dimension was not significant ($p = 1$). Thus, the short vowel /o/ had the same height of its long cognate, but it was more centralised, showing the same pattern of the previous short vowels.

4.5 Analysis of Mid-high Short Vowels Overlap

The quantitative and qualitative descriptive results, in section 4.3 above, showed a considerable overlap, particularly among the three front mid-high short vowels (/i/, /e/, /ə/) and the two mid-high short back vowels (/u/, /o/), which are plotted again in Figure 4-16 below. This section presents further analysis which has been conducted to examine the quantitative and qualitative differences between these overlapping vowels in order to report their phonetic status in SA from an acoustic perspective.

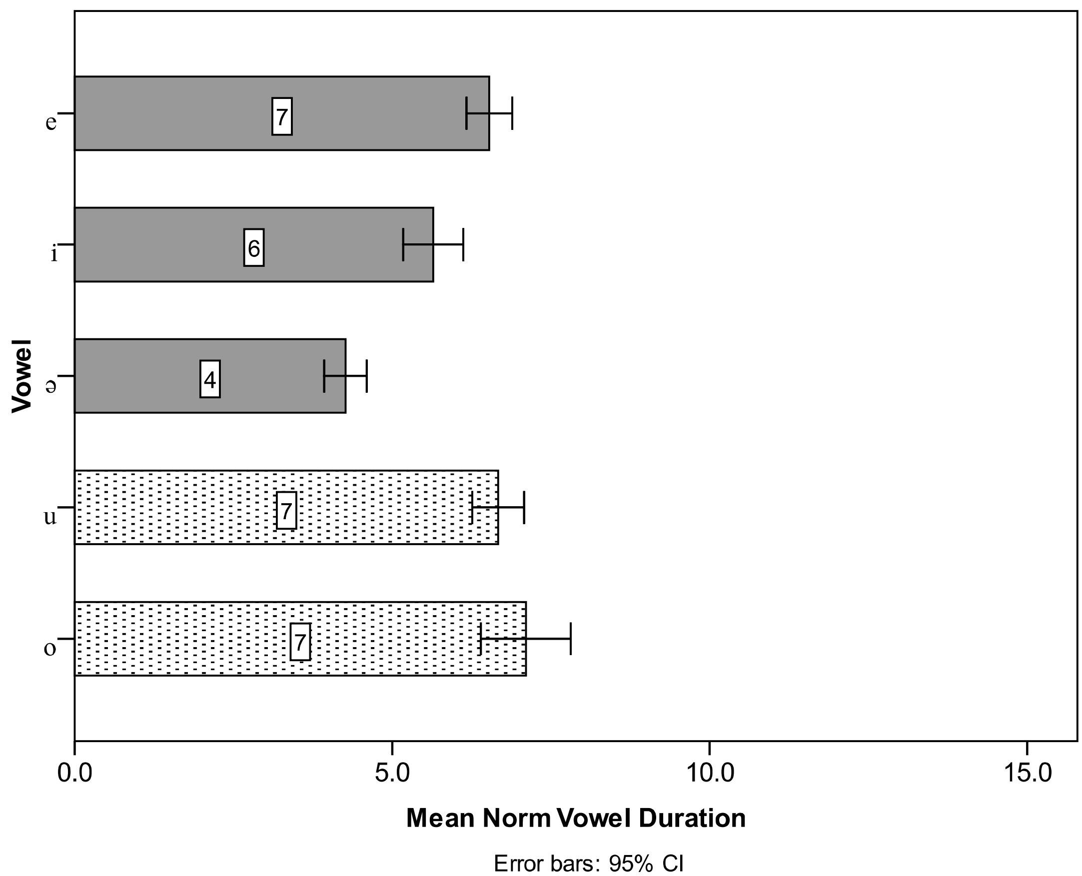

Figure 4-16: Normalised vowel duration of SA overlapping short vowels in the front mid-high area (/i/, /e/, /ə/) and in the back mid-high area (/u/, /o/)

The overlap of SA short vowels in the mid-high area was measured with respect to the Euclidean distance, which is a measure used in sociolinguistic research to investigate cases of merger (among other things) in vowel categories. The Euclidean distance calculates the distance between two points in the vowel space using the following formula (Fabricius 2007b):

$$\text{DISTANCE} = \sqrt{((\text{F1 vowel1} - \text{F1 vowel2})^2 + (\text{F2 vowel1} - \text{F2 vowel2})^2)}$$

Fabricius (2007b) suggests that the Euclidean distance can be a useful complementary measure to evaluate the separation of any two vowel categories in the vowel space. The advantage of this method is that it includes both F1 and F2 dimensions in the same measure rather than examining the difference between two vowel categories on each dimension separately.

Table 4-4: Mean Euclidean distance between overlapping SA short vowels in the mid-high area of the vowel space, and the standard deviation SD.

Vowel contrast	Euclidean Distance	SD
/i/-/e/	0.52	.40
/i/-/ə/	0.50	.34
/e/-/ə/	0.65	.37
/u/-/o/	0.65	.43

Table 4-5: Mean Euclidean distance between SA short vowels in the mid-low area of the vowel space, and the standard deviation SD.

Vowel contrast	Euclidean Distance	SD
/a/-/e/	1.53	.46
/a/-/ə/	1.85	.45
/a/-/o/	1.57	.48

Table 4-4 above presents the mean Euclidean distance of SA short vowels on the mid-high area of the vowel space. The Euclidean distances were calculated on LOBANOV normalised data points. In order to have a comparable measure to the mid-high vowels, the Euclidean distances of short vowels on the mid-low area were also calculated in Table 4-5 above. As can be seen from the tables, the mean Euclidean distances of short vowels on the mid-low area are greater than those on the mid-high area which signify the great overlap in the latter area. In particular, the Euclidean distances between the members of SA (/i/-/e/), (/i/-/ə/), and (/e/-/ə/) vowel pairs are very similar which suggests that the three SA short vowels (/i/, /e/, /ə/) might represent one vowel quality in the vowel space.

On the other hand, the Euclidean distance between SA mid-high back vowels (/u/, /o/) is slightly greater than the distance between the mid-high front short vowels. However, this distance is not as great as those between SA mid-low vowels (Table 4-5), which suggests that these two vowels might represent one vowel category as well.

A linear mixed model was used to compare the Euclidean distances of the above mid-high and mid-low vowel contrasts. The Euclidean distance was included as the dependent variable, whereas 'vowel contrast' and 'sex' were included as fixed effects, and 'participant' was used as a random effect. The results showed that 'vowel contrast' had a significant effect on the Euclidean distances of the different vowel contrasts: $F(29,1290) = 190$, $p < .000$, which means essentially that the Euclidean distances of the mid-low short vowels are significantly greater than the distances of the mid-high short vowels. Following the overall results of the present study, 'sex' had no significant effect: $F(1,1290) = .022$, $p = .882$. However, the interaction of 'vowel contrast' and 'sex' had a significant effect on the 'Euclidean distances' of the (/i/-/ə/, /a/-/ə/, /u/-/ə/, /o/-/ə/) vowel contrasts. In all of these contrasts, females had significantly greater distances than males did, which can be due to the earlier finding that male participants had significantly lower (/u/, /o/, and /ə/) and this led to shorter 'Euclidean distances' and to a narrower and less dispersed vowel space for males. Due to the physiological differences, females' vowel space has been reported in the literature to be wider and more dispersed than males' (Diehl et al. 1996, Rosner and Pickering 1994, Yang 1996).

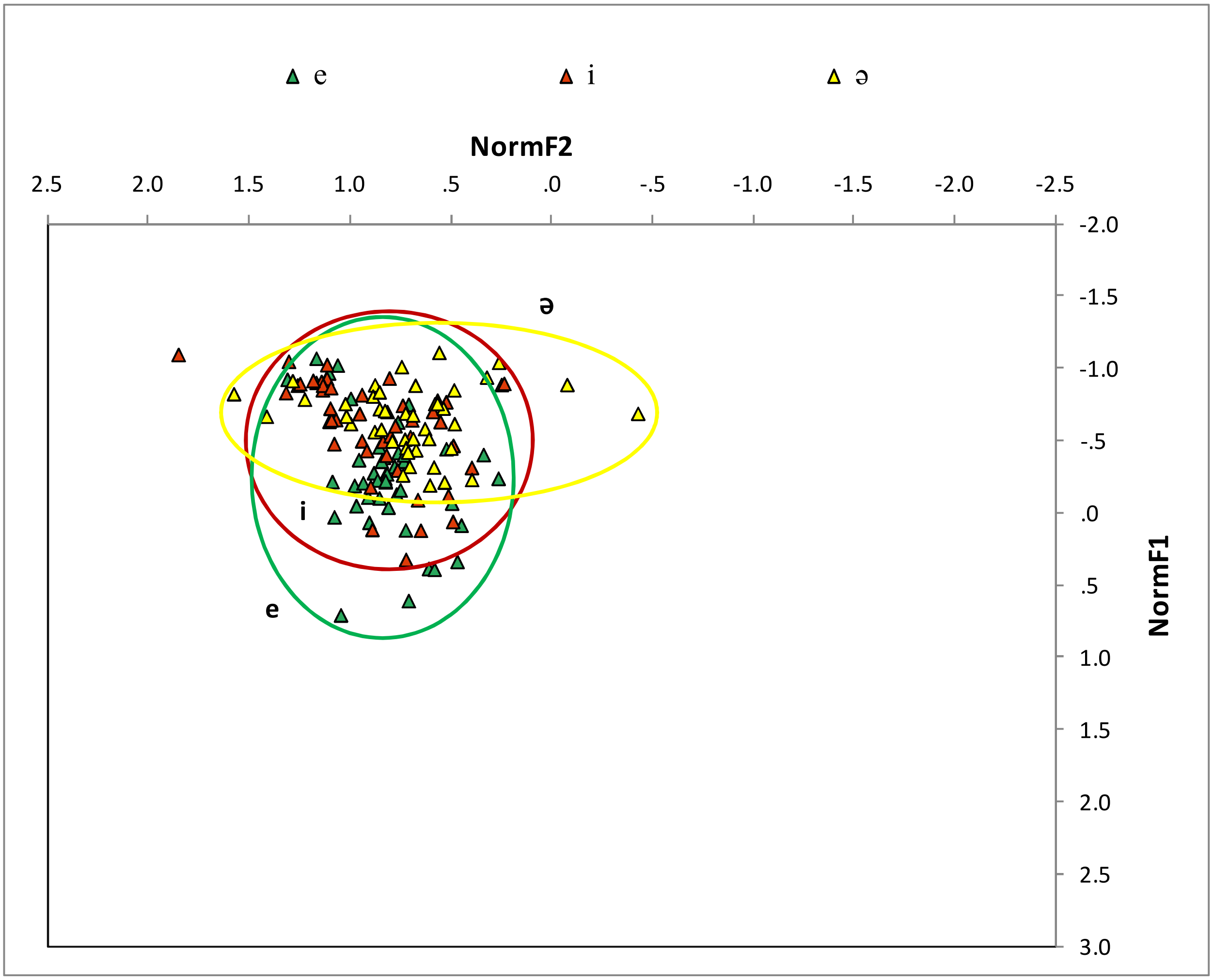

Figure 4-17: normalised F1 and F2 values of SA short vowels on the mid-high area (/i/, /e/, /ə/)

In order to determine the separation of the three SA short vowels (/i/, /e/, /ə/), shown in Figure 4-17 above, the results of the Bonferroni post hoc tests of the linear mixed model tests, conducted in section 4.3, were revisited. Table 4-6 below presents the results of the Bonferroni post hoc pairwise comparisons of SA (/i/, /e/, /ə/) vowels. According to durational differences, the results show that /i/ and /e/ are not different from each other. SA /ə/ appears to be slightly shorter than /i/, but this difference is not significant, and highly significantly shorter than the mid vowel /e/. In terms of height/F1, SA /i/ and /ə/ occupy the same mid-high area, and both are significantly higher than the mid vowel /e/. As can be seen in Table 4-6, the three vowels do not differ from each other in their F2 values. Although SA schwa seems to be slightly more centralised than /i/, the three mid-high short vowels (/i/, /e/, /ə/) occupy the same area on the front-back dimension of the vowel space.

Table 4-6: The results of Bonferroni post hoc tests for paired comparisons of SA short vowels (/i/, /e/, /ə/) in terms of vowel duration, F1, F2, and Euclidean distance[6].

vowel contrast	Norm vowel duration	Norm F1	Norm F2	vowel pairs	Euclidean distance
(/i/-/e/)	$p = 1$	$p < .000$	$p = 1$	(/i/-/e/)-(/i/-/ə/)	$p = 1$
(/i/-/ə/)	$p = .068$	$p = 1$	$p = .062$	(/i/-/e/)-(/e/-/ə/)	$p = 1$
(/e/-/ə/)	$p < .000$	$p < .000$	$p = 1$	(/i/-/ə/)-(/e/-/ə/)	$p = 1$

Similarly, the results of the Bonferroni post hoc test, examining the 'Euclidean distance' differences between the three vowel contrasts in Table 4-6, show that there are no differences in the 'Euclidean distance' between the three vowel contrasts ($p = 1$). This finding suggests that the three SA vowels (/i/, /e/, /ə/) are equally close to each other. Additionally, the 'Euclidean distances' of the above three vowel contrasts in Table 4-6 were highly significantly different ($p <.000$) from the 'Euclidean distance' of the mid-low vowel contrast (/e/-/a/), see Table 4-5 above.

[6] The results of the Bonferroni post hoc tests for some of the paired comparisons were reported as highly insignificant, i.e. p = 1. This does not mean that the vowels are identical, only that there is no evidence that they are different.

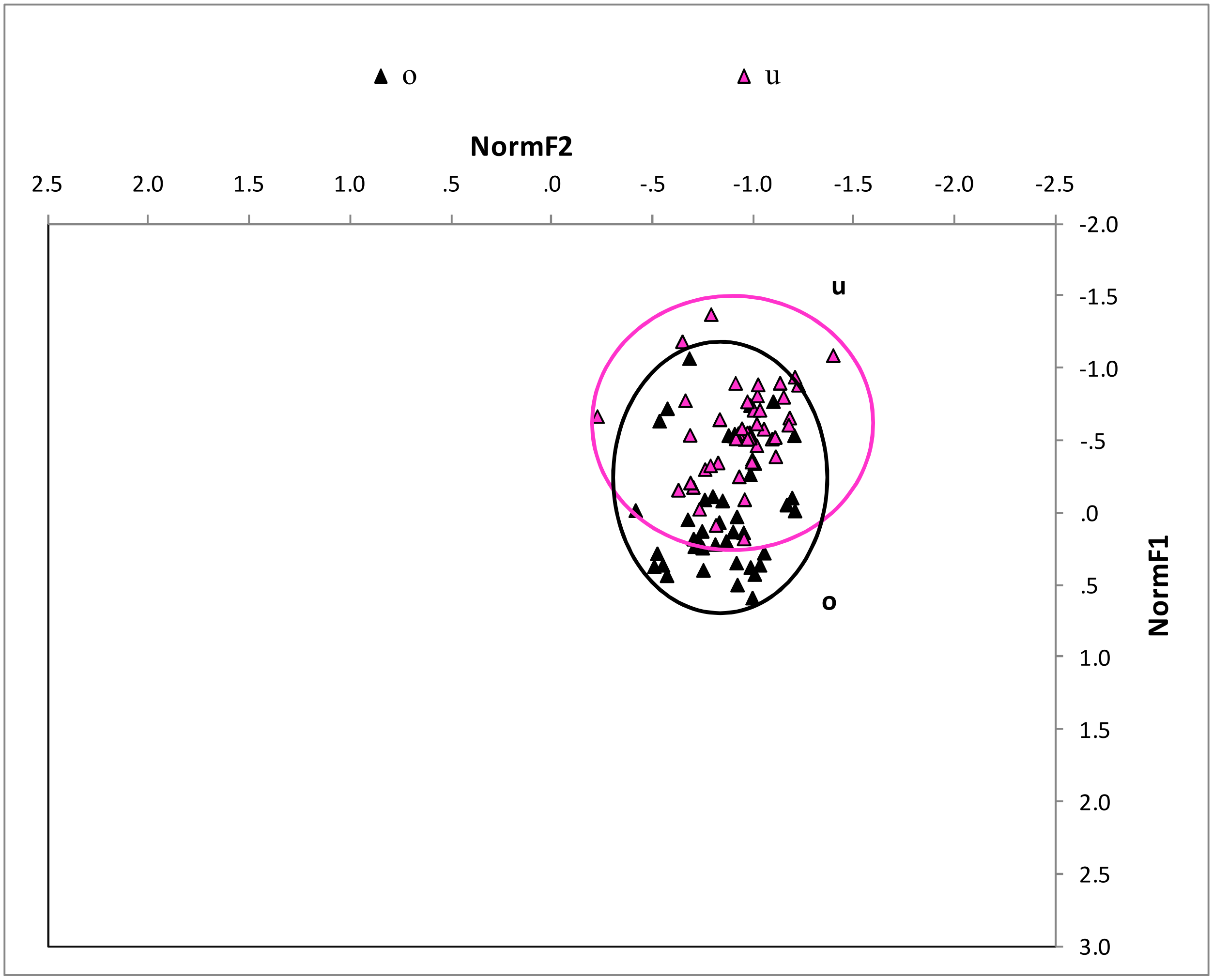

Figure 4-18: normalised F1 and F2 values of SA short vowels on the mid-high area (/u/, /o/)

As for SA short back vowels (/u/, /o/) on the mid-high area in Figure 4-18, the results of the Bonferroni post hoc tests were similar to those for SA (/i/, /e/) vowels in Table 4-6. The results showed that there was no significant difference between the short back vowels in terms of duration ($p = 1$) and F2 ($p = 1$). However, the two vowels were significantly different from each other in terms of F1 ($p < .000$), where the mid vowel /o/ was significantly lower than /u/. Additionally, the Euclidean distance between the back short vowels (0.65) was greater than the distance between the front short vowels (0.52). Nevertheless, the results of the Bonferroni post hoc test showed that the two vowel contrasts (/i/-/e/) and (/u/-/o/) were not significantly different from each other in terms of the Euclidean distance ($p = 1$), which means that the distance between the members of the (/u/-/o/) contrast is not significantly greater than the distance between the members of the (/i/-/e/) contrast. However, the distance between the members of the (/u/-/o/) contrast was highly significantly ($p < .000$), less than the distance between the mid and low vowels (/o/-/a/).

In sum, the results of the present analysis suggest that SA has a phonetic front short mid vowel quality /e/ which is significantly lower than both (/i/-/ə/) vowels. The results also show that SA (/i/-/ə/) vowels seem to represent acoustically one vowel category. Also, SA can be argued to have a phonetic short back vowel /o/ which is significantly lower than its high cognate /u/. These results should be interpreted with caution because these vowels differ from their high cognates in one acoustic dimension. Also, visually these vowels do overlap with the high vowels /i/ and /u/. Therefore, SA can be argued to have phonetic rather than phonological short mid vowels and a schwa.

4.6 Summary of Results

The overall results of the present study showed that the SA vowel system exhibits the expected durational contrast between its five short-long vowel pairs. In addition to the quantitative difference, the short-long members of these contrasts differed in quality, where short vowels tended to be more centralised than their long counterparts. SA vowel space also kept the triangular shape of Classical Arabic with the addition of mid vowels (/eː/, /e/) and (/oː/, /o/) which were more centralised than the high (/iː/, /i/) and (/uː/, /u/) vowels.

The analysis of the overlapping short vowels in the mid-high area of the vowel space revealed that SA has short mid vowel qualities /e/ and /o/, which are acoustically close in duration and frontness/F2 to their high cognates /i/ and /u/, respectively. The short mid vowels differed from the short high vowels in terms of height/F1, where mid vowels tended to be slightly lower. In terms of the Euclidean distance, which takes into account both F1 and F2, the distance between the members of (/i/, /e/) and (/u/, /o/) was not different, however, the distance was significantly smaller than the distance between mid and low vowels such as (/e/, /a/) and (/o/, /a/), which would be expected to be the same as the distance between mid and high vowels.

It has been claimed that SA has a schwa vowel (Cowell 1964), but the acoustic analysis of the present study showed that SA schwa has the same quality as the short high vowel /i/, but with a slightly more centralised F2 and a shorter duration. Furthermore, the results of the Euclidean distance revealed that the schwa is very close to both high and mid short front vowels /i/ and /e/.

Bearing in mind the results of the acoustic analysis of the present study, the following questions will be discussed:

- Since there is a quantitative and a qualitative difference between short-long vowel pairs, how do these differences compare to previous studies on SA and other Arabic dialects?
- The analysis of the overlapping short vowels showed that (/i/, /e/, /ə/) vowel categories are close to each other in quantity and quality. Similarly, (/u/, /o/) vowel categories are close to each other. However, these categories differ from each other statistically at least on one acoustic dimension. The question for these overlapping categories would be: are these phonemic or merely phonetic categories?
- If the differences between the mid-high overlapping short vowel categories are phonetic and not phonemic, will these categories be represented phonemically similar to the two phonemically established Arabic short vowels /i/ and /u/, respectively?
- What is the status of short-long mid vowel pairs? Does SA have phonemically established mid long vowels without short counterparts?
- What is the phonetic and phonological status of the schwa in SA?
- Finally, what are the implications of the results of the acoustic analysis of SA vowel system on L2 studies?

4.7 Discussion

4.7.1 SA short-long Vowel Pairs

The previous sections presented a descriptive and a statistical analysis of SA short-long vowel contrasts, in order to evaluate the quantitative and qualitative relationship between these short-long vowel pairs. A significant difference in vowel duration has been found between the members of each pair. This finding conforms to the results of Allatif et. al (2004) in which he found a significant quantitative difference between the members of SA (/uː/-/u/) and (/aː/-/a/) vowel pairs. However, Allatif et al's unpredicted finding of the lack of significant quantitative difference between SA short-long high front vowels (/iː/-/i/) was contradicted by the results of the present study. Generally, such a finding was expected for all the short-long vowel pairs, including mid vowels, since duration has a major phonological role in Modern Standard Arabic (Ryding 2005: 25).

The overall ratio of short to long vowels of the present study (1:1.63) corroborates the ratio which was calculated for Lebanese Arabic spontaneous speech (1:1.58) (Khattab 2007). However, these ratios were found to be less than the ratio calculated for a Lebanese Arabic word list (1:2.08). A similar ratio (1:2.01) was found for the Lebanese Arabic word list in Khattab et al. (2008). These results suggest that in word list tasks participants tend to exaggerate the durational difference of short/long vowels in order to give adequate clues to the identity of the vowel to avoid misidentifications. In spontaneous speech, on the other hand, speech context provides enough clues to correctly identify the vowel without relying mainly on the durational difference.

The patterns of the durational differences of short/long vowels were found to apply for singleton/geminated consonants. The ratio of singleton to geminated consonants was also reported to be greater in spontaneous speech (1:1.82) than in a word list task (1:2.29) (Khattab 2007). Additionally, the ratios reported for V to VV of the present study and for Lebanese Arabic were less than the ratios reported for C to CC (Khattab 2007, Khattab and Al-Tamimi 2008). A possible explanation might be that vowels are affected by coarticulation more than consonant.

As for quality, the members of the SA High (/iː/-/i/, /uː/-/u/) and low (/aː/-/a/) short-long vowel pairs were found to differ, significantly, from each other in both F1 (height) and F2 (frontness). On the other hand, the members of the mid vowel pairs (/eː/-/e/) and (/oː/-/o/) differed significantly from each other in F2 but not in F1. Generally, the members of SA short-long vowel pairs differed from each other in quantity as well as quality, where short vowels tended to be shorter and more centralised than their long counterparts. These results match Cowell's (1964: 9) description and Allatif's (2008) findings that SA high short vowels are slightly lower and less tense than their long counterparts. Even though a qualitative difference has been found between short-long vowel pairs, the phonological role of this difference in the phonemic distinction of these vowels cannot be determined without examining the effect of vowel quality alternations on the perception of short-long vowels (which is beyond the scope of this study).

4.7.2 SA Short Vowels

The quantitative and qualitative descriptive results of SA short vowels showed a considerable overlap between the mid-front and high-front short vowels (/i/-/e/); and the mid-back and high-back short vowels (/u/-/o/). The Euclidean distances between the members of these contrasts were significantly lower than the distances between the mid and low short vowels (/e/-/a/) and (/o/-/a/). However, the results of the acoustic analysis showed that SA has separate phonetic short mid vowels (/e/, /o/) which were significantly lower than their high counterparts (/i/, /u/). Nevertheless, the visual inspection of these categories shows a great quantitative and qualitative overlap. Additionally, the acoustic differences between mid and high short vowels are small and mainly on one dimension only, i.e. which is height/F1, which leads to question the phonological role of this difference on the phonemic status of the short mid vowels.

An example of an allophonic category in Arabic is the pharyngealised [aːˁ] category which differs phonetically from its plain counterpart /aː/. SA and most of the Arabic dialects have this pharyngealised phonetic vowel category [aːˁ] which occurs after pharyngealised segments. The pharyngealised [aːˁ] has been found to have a significantly lower F2 value than its plain cognate /aː/ (Almbark 2008, Khattab, et al. 2006). The visual inspection of (/aː/, [aːˁ]) categories in section 4.3 suggests that the acoustic difference between these categories is greater than the acoustic difference between mid and high short vowels. Nevertheless, the acoustic difference between the pharyngealised vowel and its plain cognate has been analysed and perceived by researchers and Arabic native speakers as merely phonetic. The great F2 lowering of the pharyngealised [aːˤ] does not play any phonemic role in the (/aː/, [aːˤ]) distinction in Arabic. Thus, the pharyngealised [aːˤ] has been analysed as a phonetic allophone of Arabic /aː/. In order to confirm the acoustic difference between the pharyngealised [aːˤ] and its plain cognate /aː/, the pharyngealised [aːˤ] was included in the acoustic analysis of the present study. The word /tˤaːˤr/ 'flew' was used in the analysis and the differences were examined in terms of vowel duration, F1, F2, and Euclidean distance. The results of the Bonferroni post hoc tests are presented in Table 4-7 below which shows that the acoustic differences between the plain-pharyngealised vowel pair are more prominent than the differences between the members of the two mid-high short vowel pairs. The results shows that SA (/aː/-[aːˤ]) vowels differ from each other in terms of vowel duration, F1, and F2, in which the pharyngealised [aːˤ] had a significantly longer duration, and a higher and backer vowel quality than its plain cognate /aː/.

Table 4-7: The results of the Bonferroni post hoc tests for paired comparisons of SA mid-high vowel pairs and the plain-pharyngealised vowel pair

Vowel contrast	Norm duration	F1	F2	vowel pairs	Euclidean distance
(/i/-/e/)	$p = 1$	$p < .000$	$p = 1$	(/i/-/e/)-(/u/-/o/)	$p = 1$
(/u/-/o/)	$p = 1$	$p < .000$	$p = 1$	(/i/-/e/)-(/aː/-[aːˤ])	$p = .163$
(/aː/-[aːˤ])	$P < .001$	$P < .001$	$P < .001$	(/u/-/o/)-(/aː/-[aːˤ])	$p = 1$

In terms of the Euclidean distance, the results showed that the Euclidean distance between the plain /aː/ and the pharyngealised [aːˤ] (.82) was significantly greater than the distances of the short mid-high front vowels (.52) and the short mid-high back vowels (.65). Thus, the difference between the mid and high short vowels is not great compared to the difference between the plain and the pharyngealised vowels.

Therefore, one can argue that the difference between the mid and high short vowels is mainly phonetic, similar to the difference between the plain vowel /aː/ and its pharyngealised allophone [aːˤ].

The attempt which was made to find minimal pairs to represent the suggested phonemic short mid-high vowel contrasts in section 4.1 above might support this argument. Therefore, the acoustic differences between mid and high short vowels are merely phonetic, similar to the difference between (/aː/-[aːˤ]) vowels, and do not have any phonological role in a phonemic distinction of mid-high vowel pairs. However, this argument cannot be conclusive without further phonological analysis of the above vowel contrasts, which is beyond the scope of the present study.

This argument does however coincide with Youssef's (2010) phonological analysis of mid vowels in Cairene Arabic. Youssef argues that Cairene Arabic long mid vowels have as underlying forms the Classical Arabic diphthongs /aj/ and /aw/. Thus, Classical Arabic diphthongs are claimed to be the source for dialectal Arabic mid long vowels as a result of vowel-glide assimilation process, as in Table 4-8 below. This assumption leads to the question: what is the source of dialectal Arabic mid short vowels? Youssef's suggestion for this question was that Cairene Arabic does not have mid short vowels because these vowels do not have any morphological function similar to that of the mid long vowels. For example, the (/eː/-/iː/) contrast distinguishes the masculine plural ending -iː as in [ʃatˤriːn] 'very good-plural' from the dual ending -eː as in [waladeːn] 'two boys'. Similarly, the front long vowels (/eː/-/iː/) in SA do share this morphological function with Cairene Arabic. Additionally, the SA mid back vowel contrast (/oː/-/uː/) distinguishes the masculine singular imperative form in [duːr] 'you turn' from the noun form in [doːr] 'queue'. According to my knowledge, SA mid short vowels do not have a similar morphological function to that of the mid long vowels.

Table 4-8: Examples of Classical Arabic diphthongs /aj/ and /aw/ and their dialectal forms /eː/ and /oː/ in Cairene Arabic (Youssef 2010)

Classical Arabic	Cairene Arabic	Gloss
[sˤɑjf]	[sˤeːf]	summer
[lawn]	[loːn]	colour

4.7.3 SA Schwa

Additionally, the descriptive analysis of the SA schwa (section 4.3), which is a category suggested to exist in SA by Cowell (1964: 10), shows that the schwa was found to be the shortest in duration among SA short vowels. However, the quality of the schwa was found to overlap significantly with the short SA high front vowel /i/. Surprisingly, the visual inspection of the schwa showed that it had a front quality close to /i/ vowel rather than a more centralised quality. This might be due to stressing the schwa vowel which was extracted from the bi-syllabic word /hədne/ 'a truce'; and adding a stress to the schwa gave it a more peripheral value close to SA /i/ vowel.

A closer examination of Cowell's (1964: 10) transcriptions of some of the words with a schwa vowel in Table 4-9 below reveals that most of these words would be transcribed by native speakers with /i/ vowel instead. The same words would be transcribed in Classical Arabic with one of the three classical short vowels (/i/, /u/, or /a/). Also, the word used in the present study to extract schwa productions can be transcribed as [hidne] by some native speakers, and as [hudna] in Classical Arabic. Thus, Cowell's transcriptions of the schwa can be linked historically to one of the Classical Arabic short vowels, and for native speakers all these categories merge in one vowel, i.e. /i/.

Table 4-9: Examples of SA words transcribed with a schwa vowel by Cowell (1964: 10) and their equivalents as transcribed by the author (SA native speaker) and Classical Arabic

Cowell's transcription	Native transcription	Classical Arabic	Gloss
[təlmiːz]	[tilmiːz]	[tilmiːð]	'a student (m)'
[təlᵊt]	[tilit]	[θuluθ]	'third'
[tənsa]	[tinsa]	[tansa]	'you forget'

Cowell (1964: 10) reported that the quality of the schwa vowel relies on the quality of the following consonant. For example, the schwa is argued to have a fronter quality next to plain dentals (/t/, /d/, /s/, /l/), and a lower quality next to pharyngeals (/ʕ/, /ħ/). Cowell also suggested that the schwa might have a backer quality similar to /u/

when it is next to emphatic segments (/tˤ/, /dˤ/, /sˤ/, /zˤ/), back consonants (/ɣ/, /x/) and rounding as in /w/. In accordance with Cowell's description, the results of the present study suggest that SA schwa vowel has surfaced with a higher and a more peripheral quality /i/ since the schwa was followed by the coronal stop /d/. Thus, arguably, SA has a phonetic schwa vowel category which is shorter than the other short vowels, and might surface as other short vowels depending on the quality of the following consonant or the surrounding context.

In most languages, describing and analysing the schwa is not as straightforward as it is for other vowels, due to its very short duration, which makes it affected by coarticulation more than any other vowel (Rosner and Pickering 1994). For example, Gairdner (1925: 36) described the schwa in Arabic as an allophone which substitutes for other vowels in rapid speech. Also, Cowell (1964: 10) argued that the schwa can surface as other short vowels depending on the surrounding context. In languages such as English, the schwa is associated with unstressed syllables (Roach 1991: 116). All in all, any description of the schwa in Arabic assumes other underlying forms linked to Classical Arabic, or surfaced forms linked to the surrounding context. In SA, for example, the underlying form of the schwa can be any other short vowel as suggested by Cowell (1964: 10). In the present study, the underlying Classical form of the schwa in SA was /u/ and the surfaced form was likely to be the high-front short vowel /i/, which surfaced due to the reasons mentioned earlier.

To summarise, the analysis of the acoustic results of SA vowel system shows that SA has five long vowels (/iː/, /eː/, /aː/, /oː/, /uː/), three of which have short contrasts (/i/, /a/, /u/). These short-long vowel contrasts differ from each other in quantity and in quality. The analysis also suggests that the status of SA mid short vowels /e/ and /o/ are interpreted as phonetic variants of SA high short vowels /i/ and /u/. Similarly, the status of SA schwa can be interpreted as a phonetic variant category which can surface as any other SA short vowel.

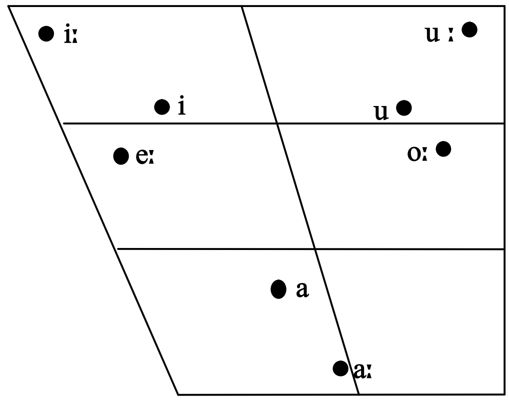

Figure 4-19: SA phonological vowel system based on the analysis of the results of the present study

In conclusion, my argument for SA phonemic vowel system is that SA has a five long and three short quality vowel system, which is represented in Figure 4-19 above. The five long and three short vowel system coincides with the descriptions and analyses of other Arabic dialects such as Gulf Arabic and Cairene Arabic, (Munro 1993, Youssef 2010). Additionally, the analysis of the results showed that SA can be argued to have four phonetic vowel categories ([e], [o], [aːˤ]) and [ə], and these allophonic categories alternate with the SA phonological vowel categories (/i/, /u/, /aː/), respectively, whereas the schwa alternates with any short vowel depending on the surrounding context. The appearance of these phonetic categories is argued to be predictable from the surrounding context. However, further phonological analysis is needed in order to determine the appearance of these phonetic categories in various consonantal contexts compared to their phonological cognates, and to determine whether they alternate regularly. Based on the analysis of the present study, Figure 4-20 below presents SA suggested phonological categories in black and the allophonic phonetic categories in grey.

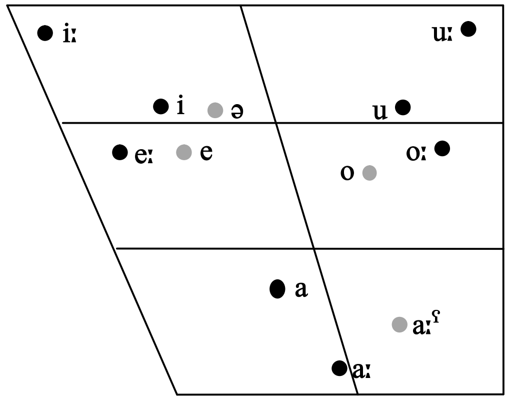

Figure 4-20: SA phonological vowel categories in black and the allophonic categories in grey based on the analysis of the present study

4.8 The Implications for L2 studies

The motivation for the acoustic analysis of SA vowel system of this chapter was to have a full analysis for the participants' first language/L1 vowel system. Therefore, it is important to state what implications the results of this analysis have on the overall aim of the present thesis, which is examining the perception and the production of English vowels as a second language/L2.

According to the results of the present study, SA has high and low short-long vowel pairs which are quantitatively and qualitatively different. This finding suggests that SA participants will not have difficulty in perceiving or producing similar English vowel contrasts. However, the phonological role for vowel length has been reported to be more prominent than vowel quality in these contrasts, which might lead to using length as a cue in their L2 perception and production (Mitleb 1981, Munro 1993). For example, Munro (1993) claims that Arabic speakers are more likely to associate the tense-lax distinction of English vowels with the quantitative short-long distinction of Arabic rather than the qualitative distinction used by native English speakers.

Additionally, the results showed that SA has distinct long mid vowels which have evolved from Classical Arabic diphthongs. The mid long vowels are claimed to have developed as a result of vowel-glide assimilation process (Youssef 2010). The development process of Classical Arabic diphthongs into mid long vowels might have implications for the perception and production of English diphthongs. For example, some of the English diphthongs, which do not match any Arabic vowel including the two /aj/ and /aw/ diphthongs, are predicted to map to Arabic mid long vowels. The English FACE and SQUARE diphthongs are predicted to be perceived and produced as the SA front mid long vowel /eː/ whereas the GOAT diphthong is predicted to be produced as the SA back mid long vowel /oː/. The English PRICE and MOUTH diphthongs are predicted to map directly to the Arabic original diphthongs /aj/ and /aw/, respectively, since the two parts of the English diphthongs are close to the Arabic ones.

The analysis of the results has revealed that SA has a three phonological short vowel system (/i/, /a/, /u/) which might lead to a greater difficulty in perceiving and producing a second language with a more crowded short vowel space, particularly if the SA participants will only refer to their short phonological categories. Many of the L2 short vowels might end up being assimilated to a single L1 vowel unless a new vowel category is developed. Best's (1995) PAM (Perceptual Assimilation Model) argues that learners of a second language use their L1 articulatory gestures when they perceive non-native segments, which means that the perception of the non-native segments depends on how similar they are to the native segments in terms of the articulatory constellations used in the native phonological space. According to PAM, SA participants are predicted to have great difficulty in their perception of English short vowels.

On the other hand, Flege's (1995: 239) SLM (Speech Learning Model) claims that L1 and L2 phonetic categories exist in the same phonological space and these phonetic categories are related at the allophonic level. He also argues that a new phonetic category can be established if a new phonetic category exists in the L2 which does not match any L1 phonetic category. According to Flege, learners can use not only their L1 phonological categories but also their L1 phonetic categories/allophones.

Bearing Flege's argument in mind, SA participants are predicted to use not only their L1 phonological vowel space but also their phonetic categories such as ([e], [o], [ə], and [aːˤ]). For example, SA mid short vowels, which are analysed as phonetic variants of their high counterparts, might aid in perceiving and producing some English/L2 new mid or mid-low short vowel categories such as the English DRESS and LOT vowels. Additionally, the lack of an acoustically distinct schwa vowel category in SA, whose quality relies mainly on the quality of the following consonant, is predicted to be reflected in the perception and production of English schwa, and further on the perception and production of English diphthongs having a schwa as one of their components such as GOAT, NEAR, SQUARE, and CURE. Furthermore, the pharyngealised allophone [aːˤ] is predicted to aid in perceiving and producing phonetically similar English vowels such as the BATH vowel.

Having analysed the phonetic and the phonological characteristics of the full vowel system of SA, the effect of L1/SA on the SA participants' perceptual and production patterns of English vowels will be better understood and reliably interpreted, which is the main goal of the following chapters.

5 The Perceptual Assimilation Task

The aim of the present and the following chapter is to examine SA participants' patterns of perception of the six English vowel contrasts, which were analysed in the pilot study in chapter 3, followed by an examination of the production patterns of these contrasts in chapter 7. Additionally, the effect of the participants' L1 full vowel system analysis, which was conducted in the previous chapter, on the perception and production of these English vowel contrasts, will be taken into consideration. In order to keep the SA participants in their English mode, the English tasks were run first. Thus, the experiments were run in the following order: the production task, the identification task, and finally Perceptual Assimilation Task (PAT). Before proceeding to the details of the PAT, the research questions of this task are reviewed as follows:

RQ1: How do SA participants perceptually assimilate English vowels to the five vowel qualities of their SA/L1 phonological inventory?

- How do SA participants perceptually assimilate the English tense vs. lax vowels to the long vs. short vowels of SA?
- How do SA participants perceptually assimilate English monophthongs vs. diphthongs to the vowels of SA?

RQ2: To what extent do PAM perceptual categories account for L2 participants' perceptual assimilation of L2 vowels to their L1 vowel inventory?

- How does SA participants' previous phonological knowledge affect their perception of English vowels?

The Perceptual Assimilation Task (PAT) of the pilot study (Chapter 3) examined the perceptual patterns of all English vowels by two SA participants in terms of a full list of their L1/SA vowel categories (Strange, et al. 1998). The results of the PAT were analysed in the framework of the PAM (Perceptual Assimilation Model) perceptual assimilation categories (Best 1994, Best and Tyler 2007). The PAT described in the present chapter was slightly modified in terms of reducing the number of English vowels examined and adding the SA pharyngealised allophonic vowel [aːˤ], and it was conducted with a larger number of SA participants. The perceptual patterns of the six

English vowel contrasts (TRAP, BATH), (FACE, SQUARE), (KIT, DRESS), (LOT, STRUT), (GOAT, THOUGHT) and (NEAR, FLEECE) are examined in terms of a full list of SA vowel system including relevant phonetic allophonic categories which were analysed in the previous chapter.

In the previous chapter, the SA vowel categories suggested by Cowell (1964) were acoustically analysed in a neutral /hVd/ context. The analysis of the results of Chapter 4 suggests that the phonological vowel system of SA consists of (/i(ː)/, /u(ː)/, /a(ː)/, /eː/, and /oː/). The results also reveal that the SA vowel system has four phonetic categories [e], [o], [aːˤ], and [ə] which were analysed as allophones of /i/, /u/, /aː/, respectively, while the schwa alternates with any SA short vowel. The effect of prior phonological knowledge on vowel perception has been noted (Ingram and Park 1997), as well as the effect of non contrastive phonetic knowledge of the native language (Best, et al. 2001). Thus, the present experiment was modified to examine the effect of the prior phonological and phonetic knowledge of SA participants on their perception of the six English vowel contrasts mentioned earlier. This aim was achieved by including both SA phonological and phonetic categories in the PAT. Additionally, extracting the relative "goodness of fit" ratings of the target English vowels as good or bad examples of the native SA vowel categories revealed the effect of SA/L1 phonological and phonetic vowel system on the participants' perceptual patterns.

In the following sections, the hypotheses and predictions for the perceptual patterns of these contrasts are presented. Additionally, the recording procedures, participants, stimulus material and the procedures for this experiment are presented, with reference made to the pilot study where procedures have not changed. After that, an illustration of data analysis procedures is given, followed by the results of the perceptual assimilation patterns for each English vowel contrast. The results of this study are discussed within PAM perceptual assimilation categories, in which English (TRAP, BATH) and (LOT, STRUT) contrasts are analysed as PAM TC (Two Category) assimilation type, English (FACE, SQUARE) as SC (Single Category), English (KIT, DRESS) and (GOAT, THOUGHT) as CG (Category Goodness), and English (FLEECE, NEAR) as UC (Uncategorised-Categorised) assimilation type. Finally, the results of this study are summarised with a discussion of the implications of these findings for the identification and discrimination of the English vowel contrasts under investigation.

5.1 Hypotheses

Table 5-1 below presents the six English vowel contrasts mentioned above which are presented and examined throughout this chapter in terms of their PAM assimilation type. The table shows the predicted categorisation/classification patterns of these vowels which were formulated based on the results of the pilot study and the analysis of SA vowel system in the previous chapters, as well as, the articulatory-phonetic similarities of the above English and SA vowel categories. In general, a significant difference is predicted in the categorisation of the members of the English (TC, CG, and UC) contrasts, while the members of the English SC contrast is predicted to have no significant difference in their categorisation patterns.

Table 5-1: The PAT predicted categorisation patterns of six English vowel contrasts into their predicted SA phonological and phonetic vowel equivalents

PAM assimilation type	English contrast	Categorisations into SA vowels
TC	(TRAP, BATH)	(/aː/, [aːˤ])
SC	(FACE, SQUARE)	/eː/
CG	(KIT, DRESS)	(/i/, [e])
CG	(LOT, STRUT)	([o], /a/)
CG	(GOAT, THOUGHT)	/oː/
UC	(FLEECE, NEAR)	/iː/

The analysis of the PAT of the pilot study showed that the English (TRAP, BATH) vowel contrast showed excellent identification and discrimination though they were categorised as a single L1 phonological category /aː/. This suggests that this contrast belongs to PAM TC (Two Category) rather than SC (Single Category) assimilation type. Thus, this contrast was analysed as a TC contrast, in which TRAP mapped to SA /aː/ and BATH was predicted to map to SA pharyngealised phonetic category [aːˤ], which would explain the excellent identification and discrimination. Therefore, the phonetic category [aːˤ] was included in the PAT of the main study. As a result, we expect the main study to show a significant difference between TRAP and BATH categorisation into their predicted L1 equivalents, i.e. they are predicted to map to two L1 categories.

The pilot study showed that English the (FACE, SQUARE) vowels were categorised as a single SA phonological category /eː/, which suggests that this contrast belongs to the PAM SC assimilation type. Similar to the previous contrast, the pilot study showed that the members of this contrast showed unexpectedly excellent identification and discrimination by SA participants. The excellent perception of this contrast was attributed to SA participants' ability to perceive adequate phonetic distance between the two vowels, particularly because these vowels are diphthongs. Additionally, SA participants might have extended the historical assimilation process of Classical Arabic vowel-glide sequences into dialectal Arabic mid long vowels when perceiving English diphthongs. Therefore, the number of categorisations of English (FACE, SQUARE) into SA mid-long vowels will be compared to the number of categorisations to SA vowel-glide /aj/ and /aw/ sequences in order to examine this prediction. As a result, we expect the main study to show no significant difference between FACE and SQUARE categorisation patterns, i.e. they are predicted to map to a single L1 category.

The pilot study showed that three English vowel contrasts were analysed as the PAM CG (Category Goodness) assimilation type: (KIT, DRESS), (LOT, STRUT), and (GOAT, THOUGHT). These contrasts were categorised as SA (/i/, [e]), ([o], /a/), and (/oː/), respectively. The pilot showed that the latter two English contrasts showed very good identification and discrimination as predicted by PAM. However, the English (KIT, DRESS) contrast showed poor discrimination and poor identification of the KIT vowel. The poor perception of this English contrast was attributed to the fact that these vowels were categorised as SA /i/ and its allophonic phonetic category [e], which are in fact perceived as a single L1 phonological category /i/. As a result, we expect the main study to show a significant difference between LOT and STRUT, and GOAT and THOUGHT categorisation patterns, i.e. they are predicted to map to two L1 categories. On the other hand, we expect the main study to show no significant difference between English KIT and DRESS categorisations, i.e. they are predicted to map to a single L1 phonological category /i/ (and some to its phonetic variant /e/).

Finally, the pilot study showed that the English (FLEECE, NEAR) vowel contrast was analysed as the PAM UC (Uncategorised-Categorised) assimilation type, in which the two vowels were mapped to the SA /iː/ vowel. As predicted by PAM, the pilot showed that this contrast had very good discrimination and excellent identification of the categorised vowel FLEECE. As a result, we expect the main study to show no significant difference between FLEECE and NEAR categorisations, i.e. they are predicted to map to a single L1 category.

As for the production of the six English vowel contrasts, the results of the pilot study showed that the members of these contrasts were produced with distinct acoustic parameters (DP). However, those productions did not match native English productions. An exception was the SC English contrast (FACE, SQUARE); the members of this contrast were produced with overlapping acoustic parameters (OP).

5.2 Methodology

The same ten male and ten female SA participants whose productions were analysed in the previous chapter, also participated as listeners in the PAT of the present study. All participants were speakers of Damascene variety. In order to have a homogeneous group of participants, they were chosen to have no exposure to English in an English speaking community (Strange, et al. 1998). All subjects had English formal education during school for at least 8 years, and to guarantee a good level of proficiency in English, all subjects were chosen to be university students in the English department at Damascus University or medium/high English level students at Asia Institute for Languages in Damascus. Similar to the pilot study, the participants completed a questionnaire about their language background, see Appendix A.

Table 5-2: Details of the 20 Syrian Arabic participants and their English skills

Participant	Gender	Age	Years of Formal education	Speaking	Understanding	Reading	Writing	Average score
S1	F	30	10	6	6	6	6	6
S2	F	23	12	7	6	7	7	6.7
S3	F	28	12	5	5	5	3	4.5
S4	F	38	10	7	7	7	7	7
S5	F	23	16	3	5	6	6	5
S6	F	47	8	4	6	5	5	5
S7	F	37	12	6	5	6	6	5.7
S8	F	24	12	6	6	6	5	5.7
S9	F	35	9	5	7	7	7	6.5
S10	F	21	10	6	7	7	6	6.5
S11	M	22	12	2	2	4	5	3.2
S12	M	29	23	5.5	6	6	5.5	5.7
S13	M	23	12	4	6	6	6	5.5
S14	M	19	11	5	6	4	4	4.7
S15	M	21	11	7	5	4	7	5.7
S16	M	21	11	3	5	4	4	4
S17	M	24	12	4	5	5	4	4.5
S18	M	24	13	4	4	5	4	4.3
S19	M	29	10	4	4	4	4	4
S20	M	23	11	3	5	5	5	4.5
Average		27.05	11.85	4.8	5.4	5.4	5.3	5.2

Table 5-2 above shows that the participants are 27.05 years old on average and they had English formal education for 11.85 years. Similar to the pilot study, the participants rated their own language skills abilities on a 7 point scale (1- very poor to 7- excellent). The table shows that the ratings are almost the same for the four skills with a slightly lower rating for speaking. The lower rating for speaking is predicted for classroom learners who have more opportunity to read, write and listen in a classroom than to speak. Additionally, the rating for understanding is higher than speaking which suggests that the participants will have greater difficulty in production than in perception.

The English recordings used as stimulus material for the PAT of this study are the same as those used in the pilot study, which were recorded following the same procedures, see section 3.1.2. For full details of the English stimuli, see section 3.1.4. Although a full list of English vowels was included in the experiment, only the perceptual patterns of the twelve English vowels under investigation (TRAP, BATH, FACE, SQUARE, KIT, DRESS, LOT, STRUT, GOAT, THOUGHT, FLEECE and NEAR) were analysed. The rest of the English vowels were included as distracters.

The alternative SA response categories were also similar to the ones used in the PAT of the pilot study. However, the pharyngealised allophone [aːˤ], which follows pharyngealised segments in Arabic, was added to the list because it was predicted to map to the English BATH vowel. The word /tˤaːˤr/ 'flew' which was used in the previous chapter to extract allophonic productions of the [aːˤ] vowel was therefore used to represent this vowel category in the present experiment. For a full list of the SA response categories, see section 3.1.4.

Similarly to the pilot study, the listening experiment was run using the Praat MFC experiment (Boersma and Weenink 2009). The participants were tested in a quiet computer room at the Asia Institute for Languages in Damascus. The procedures of the experiment were the same as the ones followed in the pilot PAT. The only difference was the addition of the pharyngealised allophone [aːˤ] to the SA alternative responses. Thus, the participants saw a list of 14 SA representative words (to represent 14 suggested SA vowel categories) instead of 13. Figure 5-1 below shows a screen shot of the main study PAT as presented to the participants, for full details of the procedures see section 3.1.5.

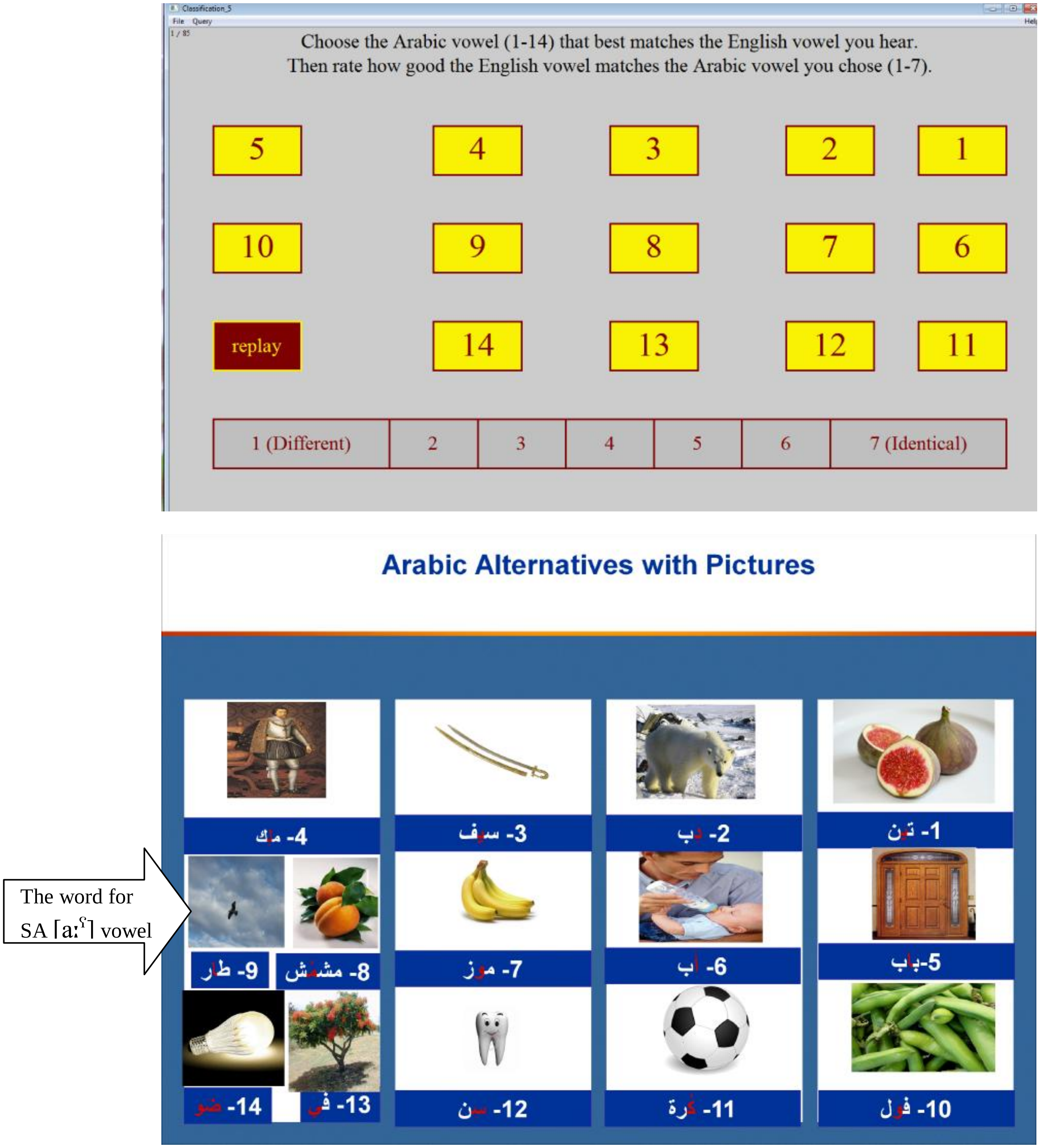

Figure 5-1: : a screen shot of the main study PAT showing praat MFC interface at the top and the pictures and written form of the 14 SA responses at the bottom

5.3 Data Analysis

Following the methods used in pilot study, a confusion matrix was created to explore the categorisation/classification of English (TRAP, and BATH, FACE, SQUARE, KIT, DRESS, LOT, STRUT, GOAT, THOUGHT, FLEECE and NEAR) vowels into a full list of SA vowels, including the pharyngealised allophone [aːˤ]. A confusion matrix was created for each SA participant; and then the categorisations for each English vowel were calculated and averaged across the participants as in Table 5-3.

Table 5-3 below presents the results averaged across 10 male and 8 female SA participants. The data for two SA females were not included for technical reasons. As shown in the table, the total number of responses for some vowels is 90, which stands for 18 (participants) x 5 (target English items) = 90 responses, whereas for other vowels it is 72, which stands for 18 (participants) x 4 (target English items) = 72.

Table 5-3: Percentages of the categorisation of six English vowel contrasts into SA vowels, showing goodness of fit ratings (in brackets: 1 different- 7 identical), and the total number of responses for each English vowels (in bold). Shaded cells show the predicted categorisation.

English vowel	/iː/	/i/	/eː/	[e]	/aː/	[aːˤ]	/a/	/oː/	[o]	/uː/	/u/	[ə]	/aj/	/aw/	NUM
TRAP		1 (.5)			52 (4.5)	17 (5)	26 (5)	1 (1.5)	1 (2.5)			2 (2)			**90**
BATH		1 (2)		4 (4.5)	56 (5)	27 (5.5)	7 (3.5)				1 (1)		1 (.5)	3 (2.5)	**90**
FACE		6 (4)	63 (5)	6 (4)		1 (2)	2 (2)				2 (2)	4 (5)	16 (5)		**90**
SQUARE	2 (2)	6 (5)	65 (5)		2 (2)	7 (4)	2 (1)			2 (2)		4 (5)	9 (6)	1 (2)	**90**
KIT	5 (4.5)	33 (5)	16 (3)	19 (4)			3 (4.5)	1 (3.5)				19 (4.5)	4 (4.5)		**72**
DRESS	2 (1.5)	17 (3.5)	27 (4)	15 (4)	2 (5.5)	1 (3)	7 (3.5)		2 (2)			24 (5)	3 (3)		**90**
LOT	1 (.5)	4 (3.5)					5 (4)	32 (4)	26 (4.5)	6 (4.5)	7 (3.5)		2 (2)	17 (3.5)	**72**
STRUT	1 (1)	13 (3.5)			17 (4.5)	4 (3)	44 (4.5)	1 (1.5)	6 (4)	1 (1.5)	4 (5.5)		2 (1.5)	7 (5)	**72**
GOAT		2 (4)				1 (1)	1 (4)	34 (4)	7 (4)	23 (5)	19 (5)	4 (2)		9 (3)	**72**
THOUGHT								68 (6)	1 (2)	13 (5)	7 (6)			11 (4)	**72**
FLEECE	80 (6)	4 (5)	7 (5)	1 (2)	1 (2)			1 (2)					6		**90**
NEAR	62 (5)		18 (5)	1 (2)		2 (4)				1 (2)		4 (2)	12 (5)		**90**

A two-sided Fisher's exact test[7] was used to examine the difference in the distribution of the responses for each English contrast into its predicted SA vowel categories, see section 5.1. For the English contrasts (FACE, SQUARE), (GOAT, THOUGHT), and (FLEECE, NEAR), which were predicted to be categorised as a single SA/L1 category, the Fisher's exact test was calculated to examine the difference in the distribution of responses between that predicted SA category and the rest of the distributions all together.

5.4 Results

5.4.1 The Perceptual Assimilation of English (TRAP, BATH) Vowels

Figure 5-2 below shows female and male categorisations of the English (TRAP, BATH) vowels into their expected responses in the participants' L1 (/aː/- [aːˤ]). As can be seen, male and female participants show similar categorisation patterns, which are averaged across all participants in Figure 5-3. More than 50% of English TRAP was categorised into its predicted SA category /aː/ and less than 20% into its pharyngealised allophone [aːˤ]. Unexpectedly, less than 30% of English BATH category was categorised into its predicted allophonic [aːˤ] category in SA and more than 50% into SA /aː/ vowel. The results of a two sided Fisher's exact test showed that there was no significant difference in the categorisation of English (TRAP, BATH) vowels into their expected responses in the participants' L1 (/aː/- [aːˤ]): ($p = 0.343$).

[7] A Fisher's exact test is a statistical significance test used in the analysis of contingency tables where sample sizes are small. Unlike the Chi square test, Fisher's exact test gives an exact P value. The tests were run using Graphpad Software GraphpadSoftware. 2012. Analyze a 2x2 contingency table. http://www.graphpad.com/quickcalcs/contingency1.cfm. April.). The actual number of responses rather than percentages was used in the test.

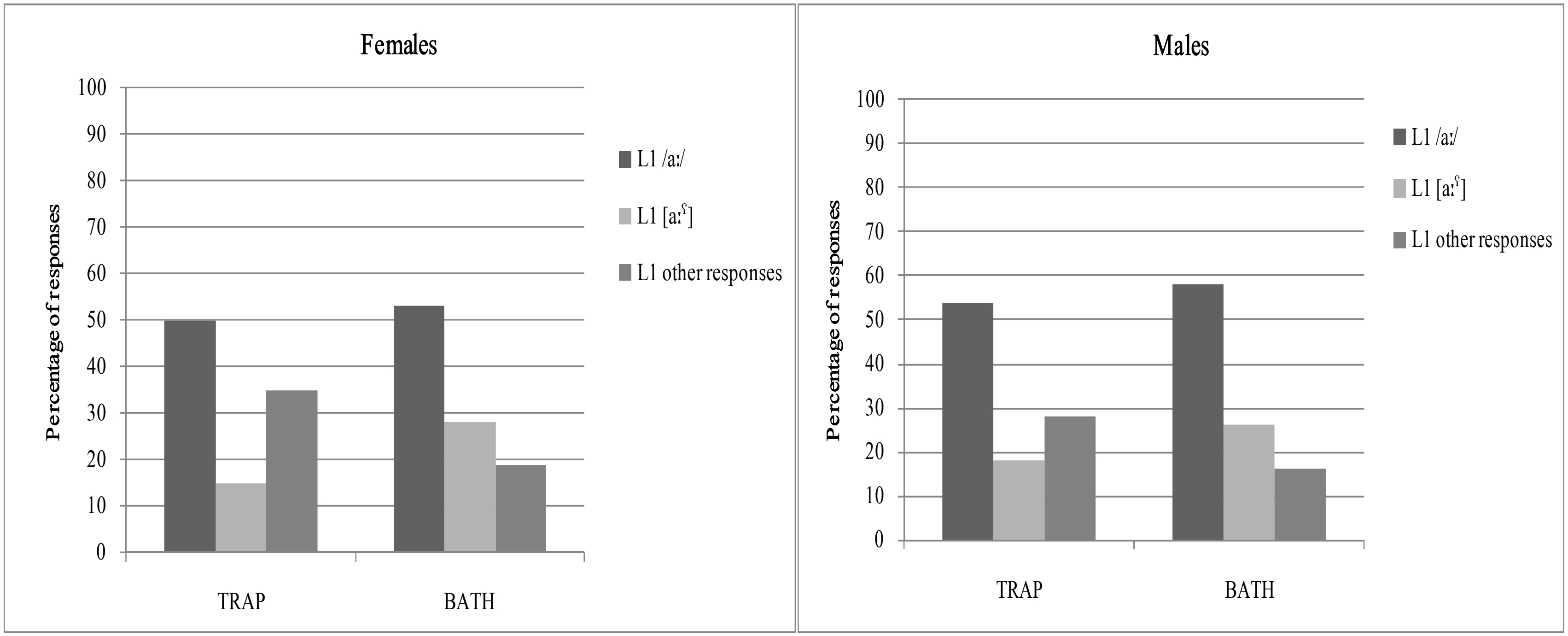

Figure 5-2: percentages of the SA female and male participants' categorisation of English (TRAP, BATH) vowels into their expected responses in SA/L1 (/aː/- /aːˤ/) (based on their articulatory-phonetic similarities and the findings of the pilot study)

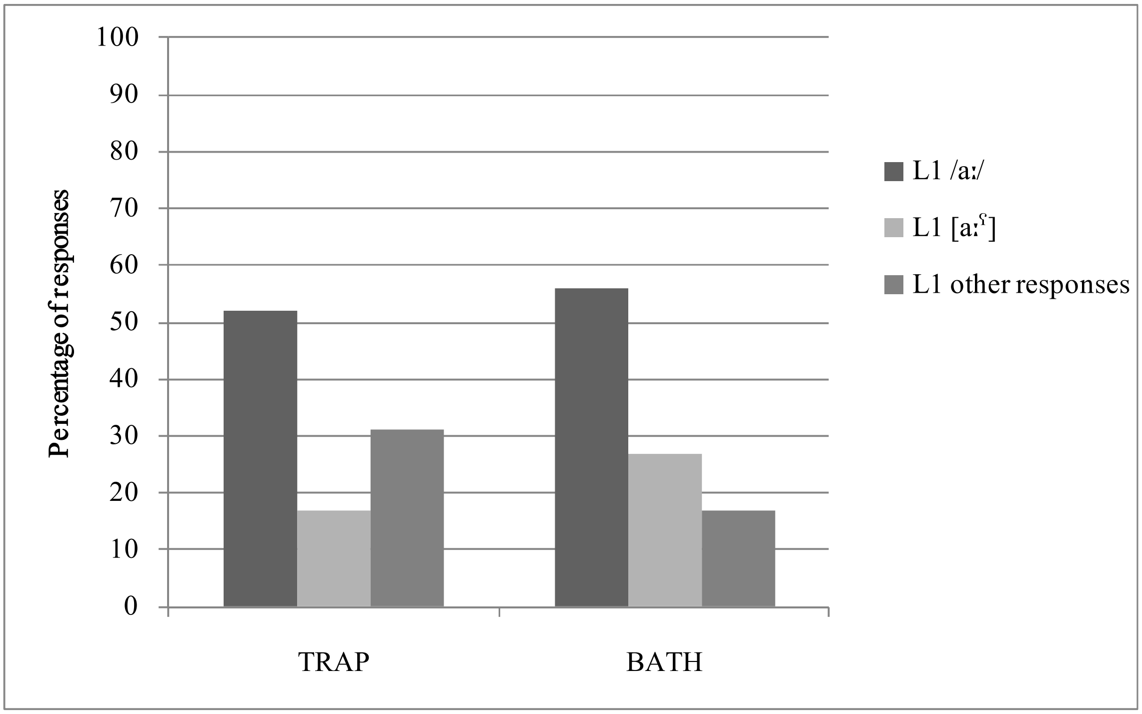

Figure 5-3: Percentages of the categorisation of English (TRAP, BATH) vowels into their expected responses in SA/L1 (/aː/- /aːˤ/) across all SA participants.

The hypothesis for the categorisation of the English (TRAP, BATH) vowels was that the members of this contrast were predicted to map to SA/L1 vowel /aː/ and its pharyngealised allophonic form [aːˤ], respectively, to form a PAM TC contrast. Thus, a significant difference was predicted to be found in the categorisation of the two vowels. The results of the PAT showed that the second part of this hypothesis was contradicted as there was no significant difference in the categorisation of the two English vowels into SA (/aː/- [aːˤ]) vowels, which means that there is a tendency to categorise the two English vowels into a single L1 category, i.e. /aː/. Additionally, the two members of the English (TRAP, BATH) contrast received high goodness of fit ratings for their similarity to their predicted L1 sounds: (4.5/7), (5.5/7), respectively. However, these ratings were not as high as expected, particularly for the TRAP vowel, which is also reflected in the percentage of categorisation. Interestingly, Table 5-3 in the previous section shows that 27% of English TRAP vowel is categorised as SA short vowel /a/. Although this percentage is not as great as the categorisation into the SA long vowel /aː/, it suggests that the English TRAP vowel might be shorter in duration than its equivalent SA long vowel, which would be expected due to the role of duration in Arabic i.e. long vowels are expected to be twice as long as their short counterparts, (this suggestion will be explored in the production experiment). For this reason, the SA participants were not giving English TRAP a very high goodness of fit rating with their long L1 vowel /aː/.

According to the results of the pilot study, this contrast was analysed as a PAM TC assimilation type, which means that the two members of this contrast mapped into two L1 vowels. In the pilot study, English TRAP mapped into SA /aː/ vowel and English BATH was predicted to map to its pharyngealised allophone [aːˤ]. However, the inclusion of the SA pharyngealised allophone [aːˤ] in the present experiment does not seem to support the prediction of the pilot since English BATH was categorised mostly as the SA /aː/ vowel. This is maybe because it is not a usual practice to ask participants to directly use their L1 phonetic allophonic categories. However, the goodness of fit rating for English BATH as the SA allophone [aːˤ] was the highest (5.5/7) which supports the suggested articulatory-phonetic similarities between these two vowels. In contrast, the analysis of the results of the present study supports the first part of the hypothesis in which English (TRAP, BATH) constitutes a TC contrast where TRAP maps to SA /aː/, in terms of the number of categorisations and goodness of fit ratings, and BATH maps into its pharyngealised allophone [aːˤ], in terms of its high goodness of fit ratings.

A possible explanation for the unexpected categorisation of English BATH and the lack of significance in the categorisation of this English contrast is that SA participants were considering the identical written forms of both SA /aː/ and [aːˤ] vowels as a single phonological category, with /aː/ as the canonical selection in their categorisations. This argument means that SA participants were not taking into consideration the phonetic allophonic categories of their L1 when they were asked to categorise English vowels. It can be argued that what they had in mind was only their phonological or phonemic vowel system, which was analysed in Chapter 4. Nonetheless, the participants are predicted to make use of their phonetic categories in their identification of these vowels in the following chapter.

5.4.2 The Perceptual Assimilation of the English (FACE, SQUARE) Vowels

The pilot study showed that the English (FACE, SQUARE) contrast represents the PAM SC assimilation type. Both English vowels were perceived equivalent to a single L1 phonological category /eː/ and they were rated as equally good exemplars of the L1 category. Thus, the hypothesis for this contrast was that it belongs to PAM SC assimilation type, and no significant difference was expected to exist between the categorisation of these two vowels into the SA predicted response /eː/ vs. the rest of the SA responses together.

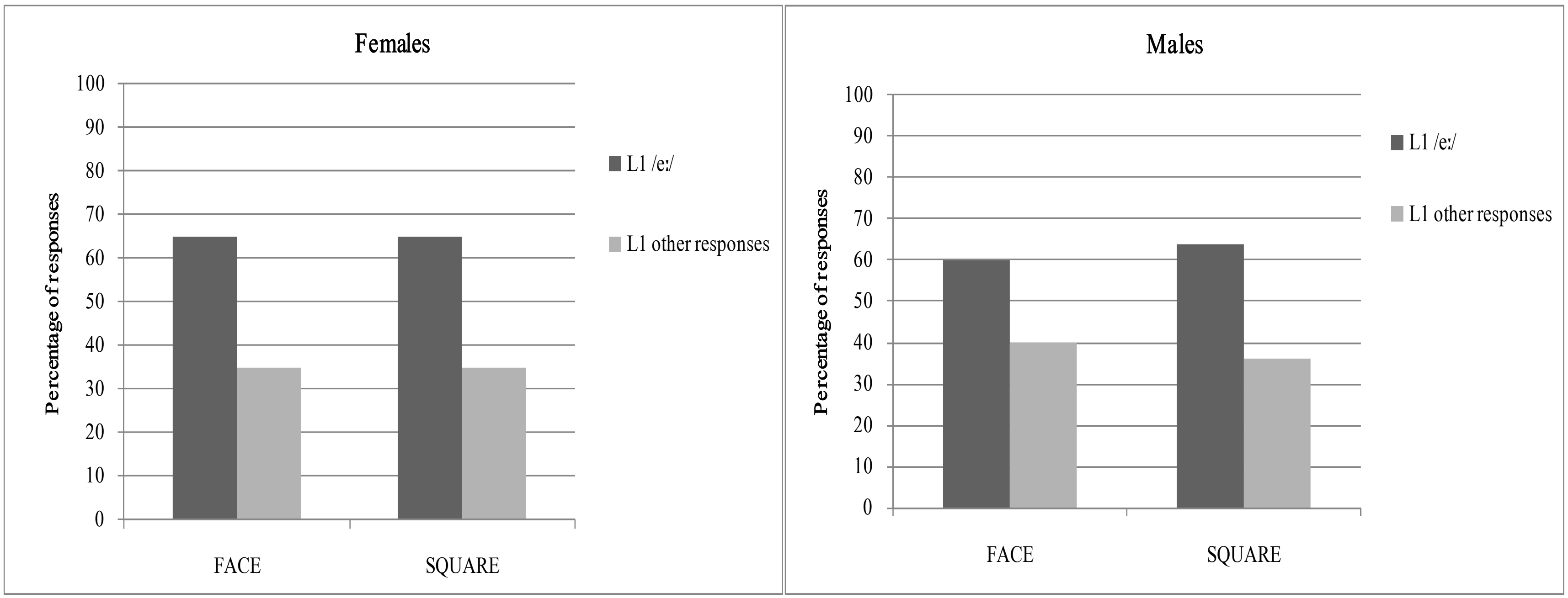

Figure 5-4: Percentages of the SA female and male participants' categorisation of English (FACE, SQUARE) vowels into their expected responses in SA/L1 /eː/ (based on their articulatory-phonetic similarities and the findings of the pilot study)

Figure 5-4 above presents the results of the present PAT for the female and male participants which shows that both have a tendency to equally categorise English FACE and SQUARE vowels into SA mid long vowel /eː/. Following the pilot study, Figure 5-5 below shows that more than 60% of the members of this English contrast were categorised as SA /eː/ with a relatively high goodness of fit rating (5/7) across all SA participants. Similar to the previous contrast, a Fisher's exact test was run to examine the difference in the categorisation of this English contrast into SA /eː/ vs. other SA responses grouped together. As predicted, the results showed there is no significant difference in the categorisation of English (FACE, SQUARE): P= .877. Therefore, this result supports the hypothesis that this is a SC contrast which maps into a single SA vowel category /eː/, with a similar goodness of fit rating 5/7.

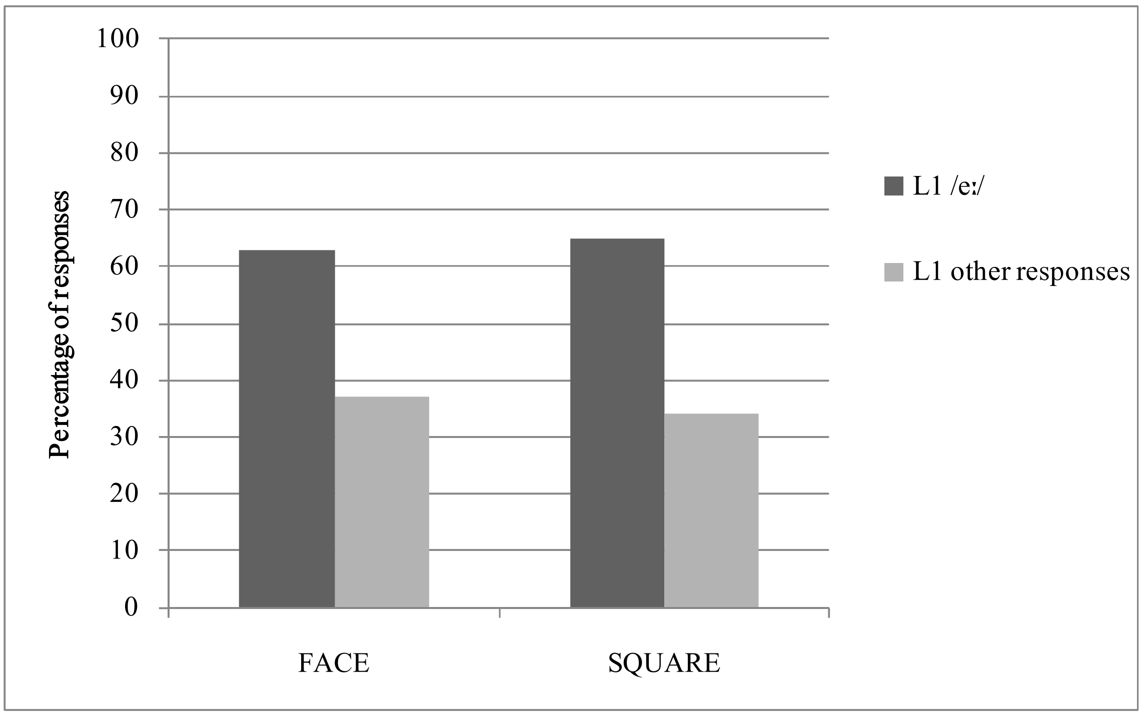

Figure 5-5: Percentages of the categorisation of English (FACE, SQUARE) vowels into their expected responses in SA/L1 /eː/ across all SA participants.

5.4.3 The Perceptual Assimilation of the English (KIT, DRESS) Vowels

The hypothesis for the English (KIT, DRESS) contrast was that it belongs to the PAM CG assimilation type, in which both map to SA phonetic category [e] with DRESS being closer to that category in terms of the number of responses and ratings. Thus, a significant difference was predicted between the categorisation of English KIT and DRESS vowels into their predicted SA vowels /i/ and [e], respectively. Figure 5-6 below presents female and male categorisations of the English (KIT, DRESS) vowel contrast. Unlike the previous contrasts, females' categorisations are slightly different from the males' with regard to the KIT vowel. As can be seen, females categorise a greater number of English KIT into its predicted SA /i/ vowel and less into its phonetic variant [e]. On the other hand, males categorise few number of KIT vowel productions into SA (/i/, [e]), and about 60% of those responses were to other SA vowel categories (/i/, /eː/, [ə]). As for the English DRESS vowel, there is little tendency to categorise it into its predicted phonetic SA category [e] by both female and male participants. Also, more than 60% of its responses were to other SA categories (/i/ 17%, /eː/ 27%, and [ə] 24%).

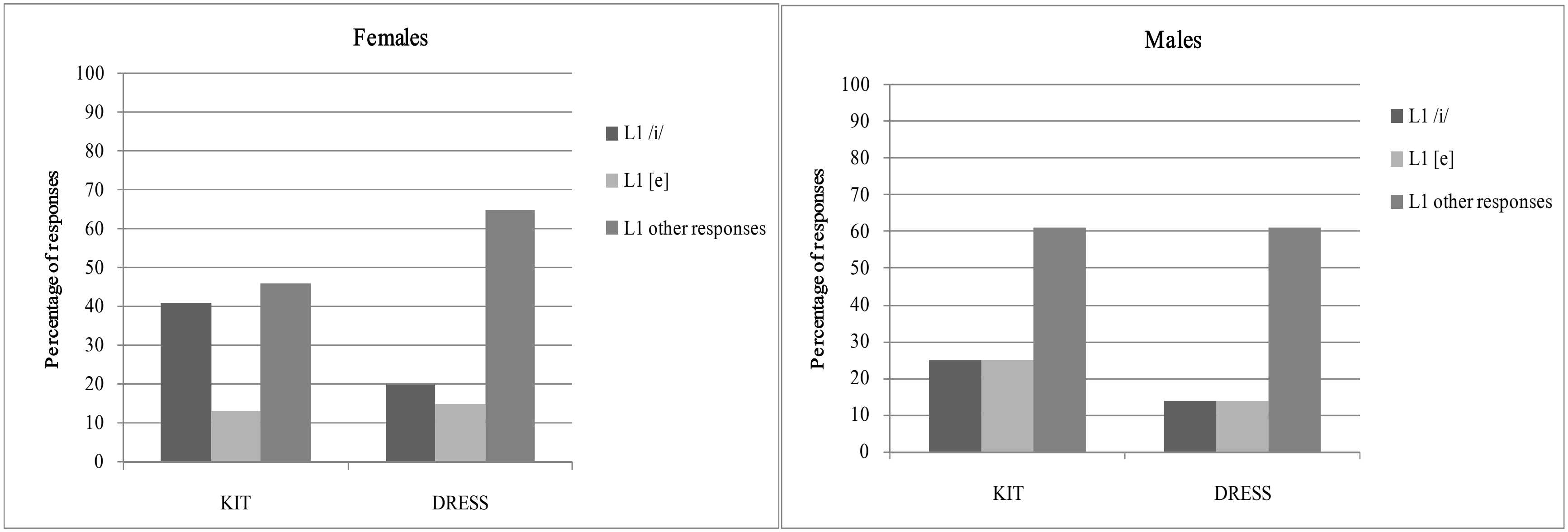

Figure 5-6: Percentages of the SA female and male participants' categorisation of English (KIT, DRESS) vowels into their expected responses in SA/L1 (/i/, [e]) (based on their articulatory-phonetic similarities and the findings of the pilot study)

Figure 5-7 below shows that across all participants there is a tendency to categorise the KIT vowel as SA /i/ or [e]. However, the percentage of those responses is not as great as expected, particularly for [e]. In particular, the percentage of the categorisation of English DRESS into its predicted category [e] represents a very small number of responses (closer to chance). The results of the Fisher's exact test showed that there was no significant difference in the categorisation of English (KIT, DRESS) into their expected responses SA (/i/, [e]) vowels: $p = 0.612$. The interpretation of this result directly means that SA participants similarly map English (KIT, DRESS) vowels to SA /i/ vowel and its phonetic category [e], which contradicts the second part of the hypothesis for this contrast.

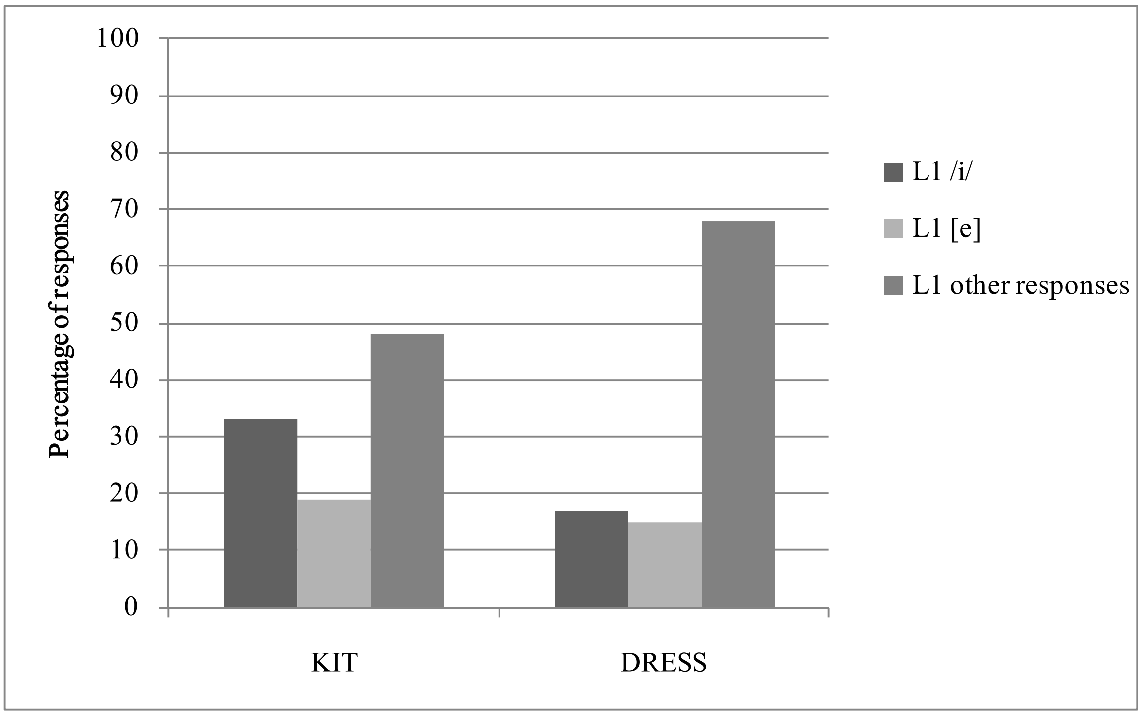

Figure 5-7: Percentages of the categorisation of the English (KIT, DRESS) vowels into their expected responses in SA/L1 (/i/, [e]) across all SA participants.

However, the result of the present study should be interpreted with caution because the number/percentage of responses mapped into the SA [e] or /i/ vowels is limited and not as great as the responses mapped to the rest of the SA vowels. Further examination of the distributions of those responses showed that the English KIT and DRESS vowels were evenly distributed into SA (/i/, [e], /eː/, and [ə]) with relatively low goodness of fit ratings. Having low ratings suggests that the English vowels share some properties with all of these vowels from an articulatory-phonetic perspective, which might be the relatively front-mid quality shared by these SA categories. Additionally, based on the argument suggested for the categorisation of the English (TRAP, BATH) vowels, SA participants might be also using their phonemic vowel inventory and not considering the phonetic categories of their language. So, analysing SA [e] and [ə] as phonetic variants of the SA vowel /i/, suggests that any categorisations into these phonetic categories should be reanalysed as categorisations into the phonological vowel category /i/ as in Figure 5-8 below.

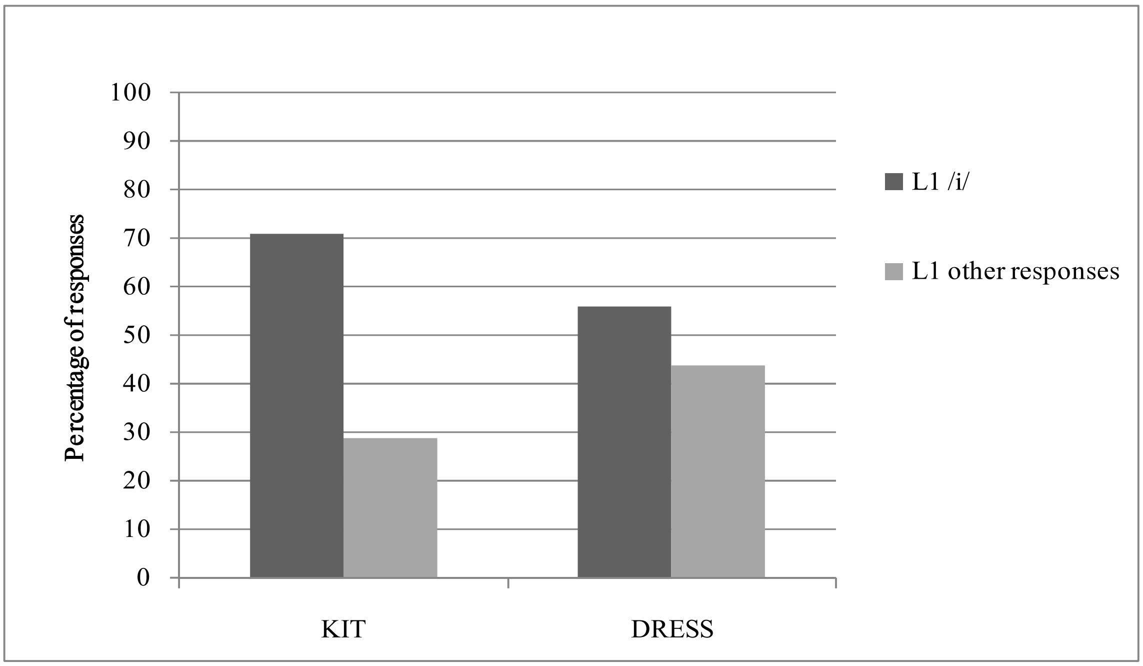

Figure 5-8: Percentages of the categorisation of the English (KIT, DRESS) vowels into the SA phonological category /i/, which includes its phonetic variants [e] and [ə])

Grouping SA phonetic categories into a single phonological category resulted in higher categorisation of both English (KIT, DRESS) vowels into that phonological category, i.e. SA /i/ vowel. Another Fisher's exact test showed that there is a significant difference in the categorisation of the English vowels into SA /i/ vs. other responses: p = 0.033, which means that the English KIT vowel was categorised as SA /i/ significantly more than English DRESS. Contrary to what was predicted, this result suggests that English KIT rather than DRESS is closer to the SA/L1 phonological category /i/.

Nevertheless, this English contrast can be analysed as PAM CG assimilation type since more than 50% of the English DRESS vowel productions were categorised as SA /i/ vowel and its phonetic variants [e] and [ə].

Therefore, the results of this analysis support the first part of the hypothesis for this contrast, i.e. it is a CG contrast. However, the English vowels map to the SA phonological category /i/ rather than its phonetic variant [e], with KIT being closer to that category in the number of responses. Also, the second part of the hypothesis is supported since a significant difference was found in the categorisations of English (KIT, DRESS) vowels, but only when grouped L1 phonological category responses were used.

5.4.4 The Perceptual Assimilation of the English (LOT, STRUT) Vowels

Based on the findings of the pilot study, the hypothesis for the categorisation of the English (LOT, STRUT) contrast was that it constitutes a PAM CG assimilation type. The two members of this contrast were predicted to map to SA/L1 vowel /a/, with STRUT being closer to that category. Thus, a significant difference was expected in the categorisation of these vowels into their predicted equivalents in SA ([o], /a/). Similar to the previous contrasts, Figure 5-9 below shows female and male categorisations of English (LOT, STRUT) vowels. As can be seen, the categorisation patterns are very similar for both, with more categorisation of LOT into its predicted SA category [o] by females.

Figure 5-10, which presents the same categorisations averaged across all participants, shows that there is a tendency to categorise English LOT and STRUT vowels into their predicted categories in SA [o] and /a/, respectively. However, the percentages of those categorisations were not as high as expected. Additionally, the English (LOT, STRUT) vowels were given a relatively similar low goodness of fit rating (4.5/7) to their predicted SA equivalents. These findings suggest that these English vowels share some articulatory-phonetic properties with the SA categories but they are not a match, which is reflected in the low goodness of fit ratings. The results of the Fisher's exact test for the categorisation differences for this contrast showed that there was a significant difference in the categorisation of English (LOT, STRUT) into SA ([o]- /a/): p= 0.0001.

Unlike the SA phonetic category [e] used in the categorisation of the previous English (KIT, DRESS) contrast, the results of the present study provide evidence that SA participants were using the phonetic category [o] in their categorisations of English LOT vowel, without referring to its phonological category /u/ (only 7%). Nevertheless, more than 60% of English LOT and 50% of STRUT were categorised as other SA vowel categories. Further examination of the data showed that the English (LOT, STRUT) vowels were mapped to SA long vowels /oː/ (32%) and /aː/ (17%), respectively, with relatively low goodness of fit ratings (4/7) and (4.5/7).

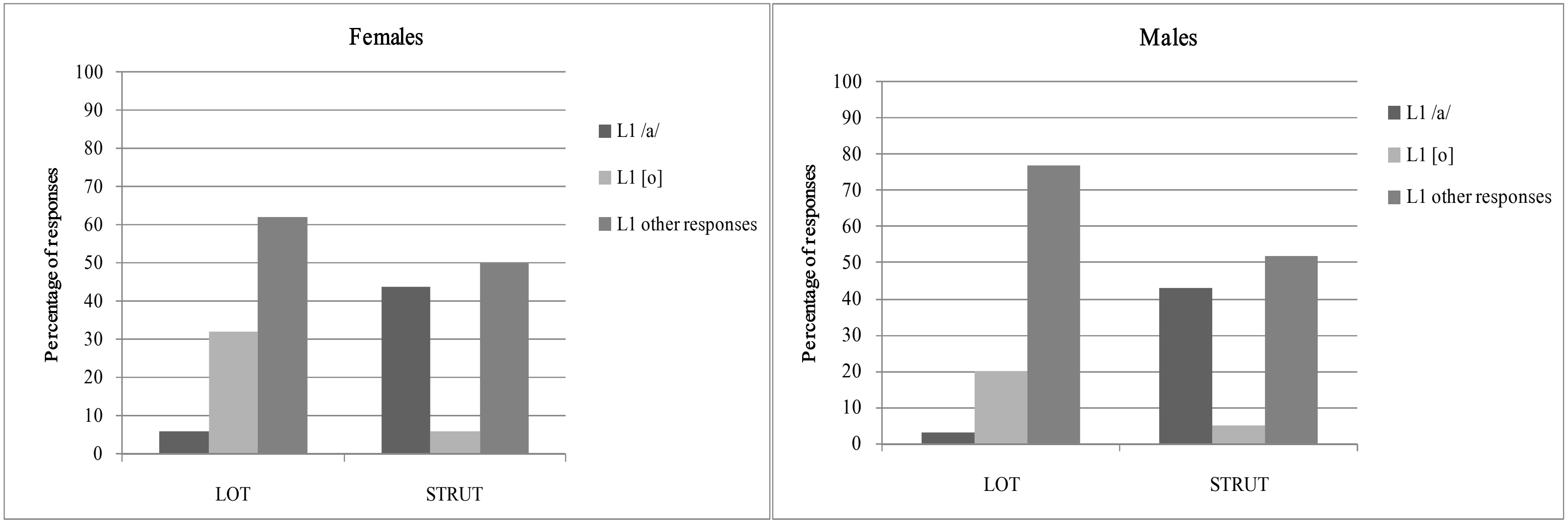

Figure 5-9: Percentages of the SA female and male participants' categorisation of the English (LOT, STRUT) vowels into their expected responses in SA/L1 ([o]- /a/) (based on their articulatory-phonetic similarities and the findings of the pilot study

Figure 5-10: Percentages of the categorisation of the English (LOT, STRUT) vowels into their expected responses in SA/L1 ([o], /a/) across all SA participants.

The results of the PAT of the present study supported the second part of the hypothesis for this contrast, as the STRUT vowel was categorised as SA /a/ significantly more often than English LOT was. However, the percentage of categorisation of English LOT as SA /a/ represents a small number of responses (5%) (which is close to chance). Therefore, the English (LOT, STRUT) contrast was found to map to two separate SA categories [o] and /a/, respectively, or to their long counter parts /oː/ and /aː/. This may be because of the role of length in Arabic, i.e. SA short vowels are shorter than their English lax equivalents to maintain a more prominent length distinction between long and short vowels. The implication of this result is that this contrast should be reanalysed as PAM TC assimilation type rather than CG since English (LOT, STRUT) vowels are mapped into two different SA categories, whether these categories are long or short.

The results for the present contrast and the previous one show that there is a tendency to categorise the English lax vowels DRESS, LOT, and STRUT into the SA long vowels /eː/, /oː/ and /aː/. This finding does not coincide with the expected categorisation of English lax vowels into SA/L1 short vowels, which suggests that the participants are using a feature other than (or in addition to) vowel duration to categorise

English vowels. The implications of this finding will be discussed further in section 5.5 below.

5.4.5 The Perceptual Assimilation of the English (GOAT, THOUGHT) Vowels

The hypothesis for this contrast was that it belongs to the PAM CG assimilation type, in which the English (GOAT, THOUGHT) vowels map to SA mid long vowel /oː/, with THOUGHT being closer to that category in the number of responses and ratings. Thus, a significant difference was predicted in the categorisation of this contrast into SA /oː/ vs. other SA categories grouped together. Figure 5-11 below shows that female and male participants show similar categorisation patterns of the English (GOAT, THOUGHT) vowels into their expected SA category /oː/. The same categorisations are presented again across all participants in Figure 5-12 which shows that English THOUGHT is categorised as SA /oː/ more than English GOAT. Additionally, English THOUGHT was given a higher goodness of fit rating (6/7) than English GOAT (4/7). This finding suggests that English THOUGHT is closer to its predicted SA equivalent vowel /oː/, from an articulatory-phonetic perspective. These findings were confirmed by the results of the Fisher's exact test which showed that English THOUGHT was categorised as SA /oː/ significantly more than English GOAT: $P = 0.0001$.

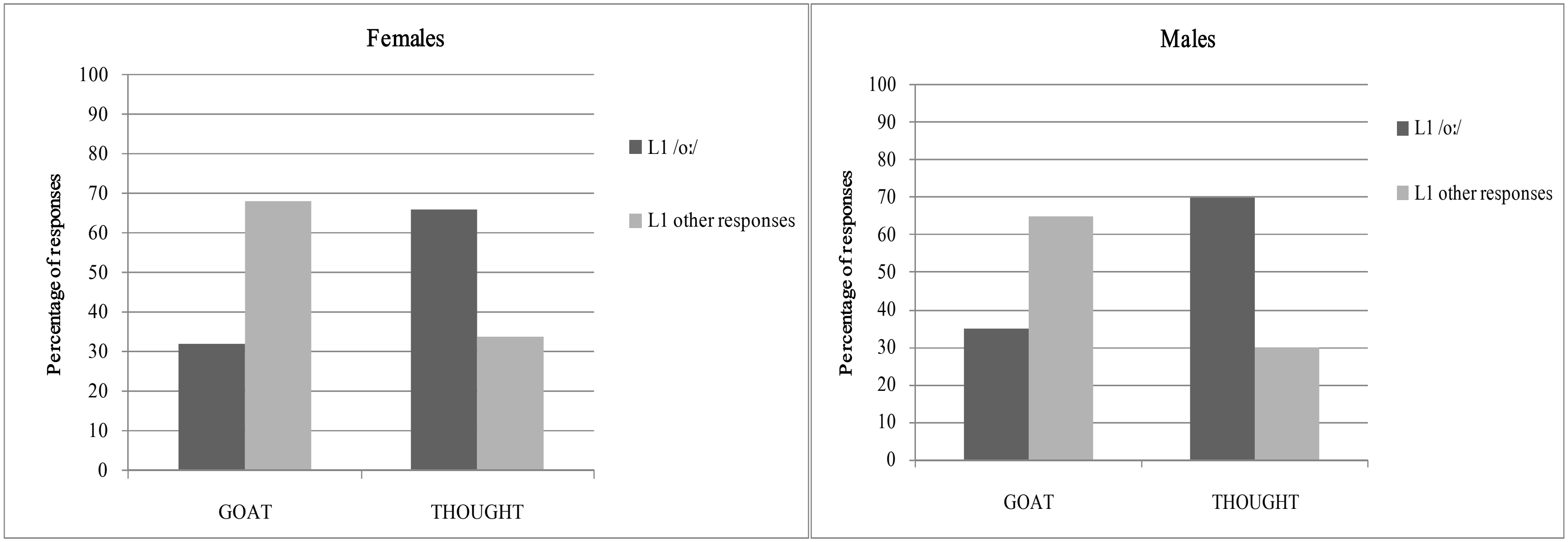

Figure 5-11: Percentages of the SA female and male participants' categorisation of the English (GOAT, THOUGHT) vowels into their expected response in SA/L1 /oː/ (based on their articulatory-phonetic similarities and the findings of the pilot study

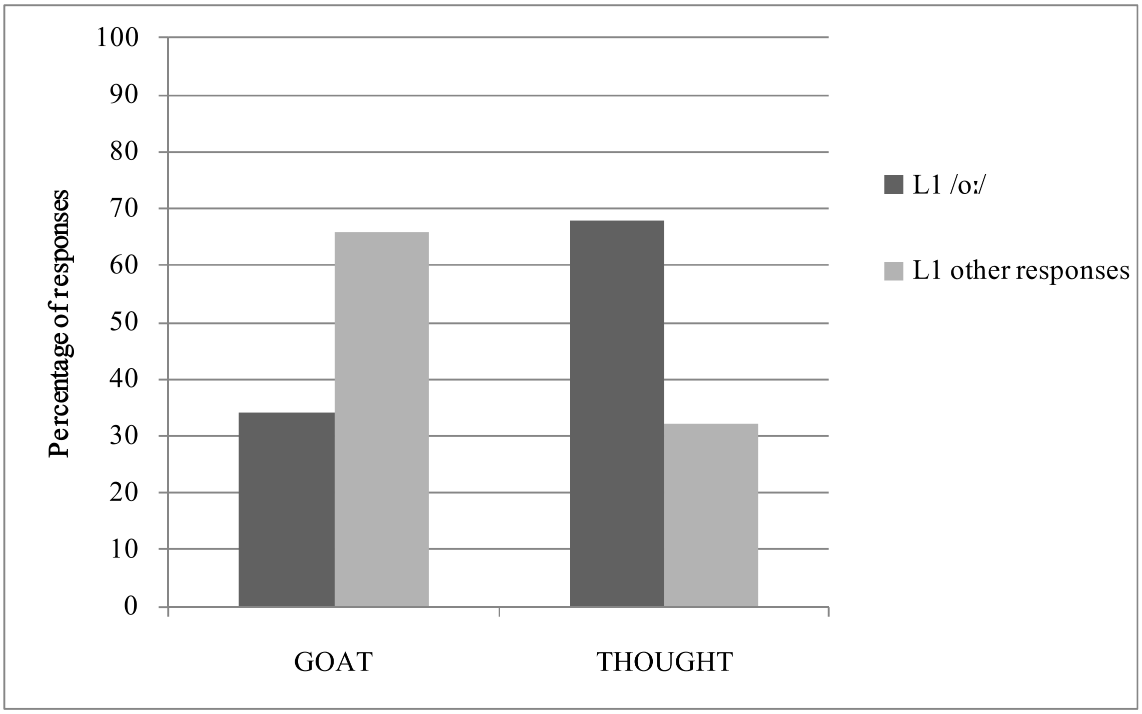

Figure 5-12: Percentages of the categorisation of the English (GOAT, THOUGHT) vowels into their expected responses in SA/L1 /oː/ across all SA participants.

On the other hand, more than 60% of English GOAT was categorised as other SA vowel categories. Further examination showed that these categorisations were mostly as SA high back vowels /uː/ (23%) and /u/ (19%), with relatively similar high goodness of fit ratings (5/7). Similarly, 13% of English THOUGHT was categorised as the SA high long back vowel /uː/ with a similar rating (5/7). The categorisation of English GOAT into three SA back vowels with similar high goodness of fit ratings suggests that it shares some articulatory-phonetic properties with all these vowels. It also suggests that GOAT is equally close to those SA categories, but not an identical match to any of them. This finding can be due to the diphthongal nature of English GOAT vowel which is not directly similar to any SA vowel. Thus, the results of the present study support the hypothesis for this contrast as English THOUGHT was closer to the SA vowel category /oː/, and it was significantly categorised as this category more than English GOAT, with a higher goodness of fit rating.

5.4.6 The Perceptual Assimilation of the English (FLEECE, NEAR) Vowels

The hypothesis for this contrast was that it represents the PAM UC assimilation type, in which both English vowels were predicted to map to the SA /iː/ vowel. English NEAR was analysed as the Uncategorised vowel, and FLEECE was the Categorised one, which had higher categorisations into the SA vowel and higher goodness of fit rating. Therefore, a significant difference was predicted between the categorisation of these English vowels into their predicted L1 equivalent /iː/ vowel vs. other L1 responses.

Unlike the previous English vowel contrasts, the English (FLEECE, NEAR) contrast received the highest percentage of categorisation into their predicted SA vowel /iː/, particularly for the English FLEECE vowel. Figure 5-13 below shows that female and male categorisations of this contrast have the same pattern with slightly higher categorisations for females. As predicted, a great number of English FLEECE productions were categorised as SA /iː/ with a high goodness of fit rating (6/7), whereas only 62% of English NEAR productions were categorised as SA /iː/, with a slightly lower goodness of fit rating (5/7), as shown in Figure 5-14. The results of the Fisher's exact test showed that English FLEECE was categorised as SA /iː/ significantly more than English NEAR: $p = 0.014$. Therefore, the results of the present study support the hypothesis for this contrast since the English (FLEECE, NEAR) contrast was categorised as SA /iː/, with FLEECE being clearly shown to be the equivalent of the SA vowel in terms of the number of responses and the ratings.

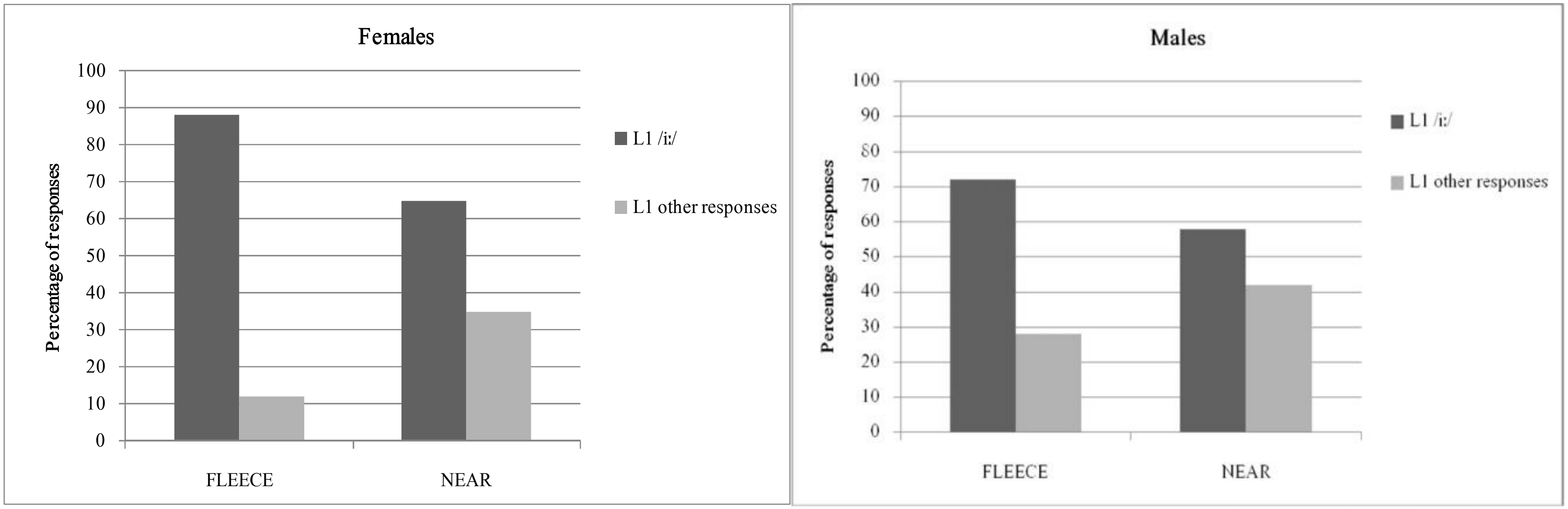

Figure 5-13: Percentages of the SA female and male participants' categorisation of the English (FLEECE, NEAR) vowels into their expected response in SA/L1 /iː/ (based on their articulatory-phonetic similarities and the findings of the pilot study

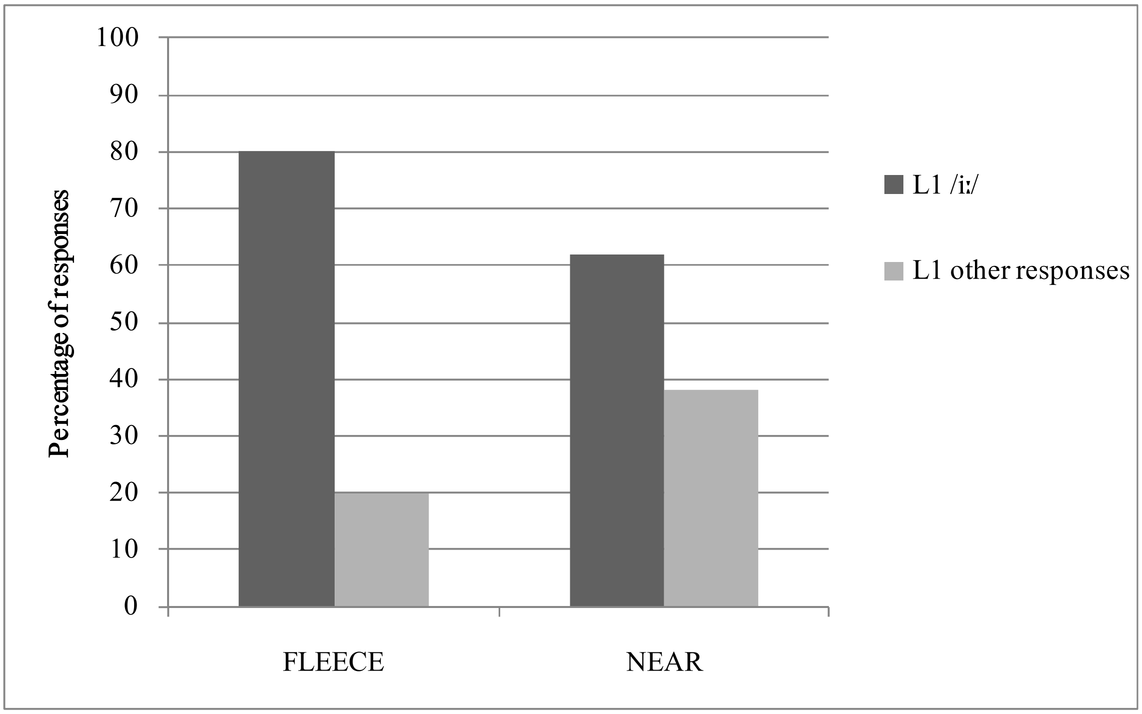

Figure 5-14: Percentages of the categorisation of the English (FLEECE, NEAR) vowels into their expected responses in SA/L1 /i:/ across all SA participants.

5.5 Discussion

The PAT of the present study aimed to examine the perceptual assimilation of six English vowel contrasts into their predicted equivalents in SA. The similarity between those English vowels and their SA equivalents was judged based on their articulatory-phonetic similarities and the results of the pilot study PAT which was conducted on a full list of English and SA vowels. Table 5-4 below presents a summary of the PAT conducted in the pilot and in the present study for the six English vowel contrasts under investigation. It is important to highlight that the PAT is a demanding task, in which the participants were asked to think in their L1 and map English vowels into their 14 L1 options. However, since these options were the same throughout the task, it can be argued that the participants managed to do the task reasonably well. The results of the present experiment were analysed using PAM perceptual assimilation categories which are discussed as follows: English TC contrasts (TRAP, BATH) and (LOT, STRUT), English SC contrast (FACE, SQUARE), English CG contrasts (KIT, DRESS) and (GOAT, THOUGHT), and English UC contrast (FLEECE, NEAR).

Table 5-4: A summary of the PAT results of the pilot and the present study for six English vowel contrasts, which were analysed within PAM framework

English Vowel Contrast	Pilot Results		PAT Results	
	SA equivalents	PAM Assimilation type	SA equivalents	PAM Assimilation type
(TRAP, BATH)	(/aː/, [aːˤ])	TC	(/a(ː)/,[aːˤ])	TC
(FACE, SQUARE)	/eː/	SC	/eː/	SC
(KIT, DRESS)	(/i/, [e])	CG	/i/	CG
(LOT, STRUT)	([o], /a/)	CG	([o](/oː/), /a/)	TC
(GOAT, THOUGHT)	/oː/	CG	/oː/	CG
(FLEECE, NEAR)	/iː/	UC	/iː/	UC

5.5.1 English Two Category (TC) Contrasts

In the pilot study, the English (TRAP, BATH) vowels were found to map to a single SA phonological category /aː/ which did not coincide with the excellent discrimination and identification of these vowels. As can be seen in Table 5-4, the English (TRAP, BATH) vowel contrast was analysed as PAM TC assimilation type, which means that these vowels were predicted to map into two SA vowel categories. From an articulatory-phonetic perspective, the closest SA equivalents were found to be the phonological category /aː/ and its pharyngealised allophone [aːˤ], respectively.

Table 5-4 above shows that the analysis of the present study partially supported the predictions of the pilot. However, no significant difference was found in the categorisations of English (TRAP, BATH) into SA (/aː/, [aːˤ]) vowels, and they were mostly categorised as the phonological category /aː/ (52% and 56%, respectively); this contrast was analysed as PAM TC assimilation type. Also, categorising 27% of English TRAP to the SA short vowel /a/ with a high goodness of fit rating 5/7 suggested that the participants perceived TRAP as shorter in duration than their L1 long vowel /aː/. This result was attributed to the role of vowel duration in Arabic, in which long vowels are one and a half as long as their short counterparts, in general. In particular, SA /aː/ vowel was almost twice as its short counterpart /a/ (1.91:1). Therefore, I predict that English TRAP should be produced twice as long as SA short /a/ to be perceived as SA long /aː/.

The high goodness of fit ratings of English BATH as SA [aːˤ] supported their predicted perceived similarities by SA participants. However, when it came to categorisation, SA participants showed a preference towards the phonological category /aː/ (or its short variant /a/) more than its phonetic variant [aːˤ]. One can argue that SA participants might have considered the two SA categories (/aː/, [aːˤ]) as one phoneme (they are indeed), and they ignored or did not realise the phonetic difference between these categories. As a result, the phonological category /aː/ was mostly chosen by the participants to represent SA /aː/ and its allophone [aːˤ]; thus, SA /aː/ was chosen to be the closest equivalent vowel to both the English (TRAP, BATH) vowels.

According to the results of the PAT of the main study, the English (LOT, STRUT) vowel contrast was also analysed as a TC contrast. This analysis contradicted what was found in the pilot study for this contrast, in which (LOT, STRUT) was analysed as the PAM CG assimilation type. The results of the main study showed that English LOT was categorised as its predicted SA equivalent [o] with a relatively low rating; and unlike the pilot the categorisation of this vowel as SA /a/ was close to chance and it was given a low rating of similarity. Unexpectedly, LOT was also categorised as SA mid long vowel /oː/ more often than its short form [o] but with a lower goodness of fit rating. This result can be attributed to perceiving English LOT as longer in duration than the SA short vowel [o]. Similar to English TRAP, the role of duration in Arabic may have had an effect on the categorisation of English LOT.

The categorisation of English STRUT, on the other hand, did not show a different pattern from the pilot as it was categorised mainly as SA /a/ but with a lower rating of similarity. Therefore, the English (LOT, STRUT) contrast was reanalysed as the PAM TC assimilation type since the two English vowels mapped into two separate SA vowel categories ([o], /a/). This analysis also coincides with very good identification and discrimination of the two English vowels in the pilot, see section 3.1.

Interestingly, the two English TC contrasts of the present study mapped into two SA phonological vs. phonetic vowel categories, i.e. (/aː/, [aːˤ]) and ([o], /a/). The difference between these two contrasts was that English (TRAP, BATH) vowels mapped into a phonologically related SA vowel pair (/aː/, [aːˤ]), whereas English (LOT, STRUT) vowels mapped into a non-phonologically related SA contrast ([o], /a/). This difference may explain why it was easier for the participants to categorise English (LOT, STRUT) vowels into their predicted categories than English (TRAP, BATH). Therefore, it can be argued that SA participants were able to use the allophonic phonetic vowel categories of their L1 to identify and discriminate these English vowel contrasts, which would have been more difficult if they relied only on their phonological vowel inventory.

These results indeed suggest that L2 learners indirectly use their phonological as well as their phonetic vowel systems when they perceive L2 vowels. Thus, it can be argued that L2 learners use their full allophonic inventory (which they might not be aware of) as possible phonetic categories in their L2 learning. This is particularly helpful for Arab learners of English because their L1 vowel system has a limited number of phonemes and adding the phonetic allophones of their L1 will give them more options and make it easier for them to learn a second language with a rich vowel system such as English.

5.5.2 English Single Category (SC) Contrast

As for the English (FACE, SQUARE) vowels, the results of the present study replicated the findings of the pilot. Both vowels were equally mapped onto the SA front mid long vowel /eː/ with similarly high goodness of fit ratings. Thus, this contrast was analysed as a PAM SC assimilation type. Although SA does not have parallel pure diphthongs to map directly to English diphthongs, it has the vowel-glide sequences /aj/ and /aw/. These diphthong-like sounds changed historically into mid long vowels (/eː/, /oː/) due to a vowel-glide phonological assimilation process (Youssef 2010). Therefore, a possible explanation for categorising two English diphthongs as a single SA mid long vowel might be extending this historical phonological process into English diphthongs, i.e. perceiving and/or producing them as mid long vowels. Additionally, the phonological process of changing a vowel-glide sequence into a mid long vowel does not mean that SA speakers do not perceive the phonetic properties of diphthongisation when they hear it. Particularly, there are many Arabic words which are still used with their original vowel-glide sequences due to their lexical status as formal words such as [raj] ‘irrigation’ and [dawla] ‘country’.

In general, PAM predicts that L2 learners in their early stages will have some difficulty discriminating the members of a SC contrast, at least till a new phonological category is established to one of the L2 sounds (Best 1995). However, it has been argued that the SA participants in the pilot study were able to perceive an adequate phonetic distance between the English FACE and SQUARE vowels. Also, the English (FACE, SQUARE) contrast, unlike the other English contrasts, consists of two diphthongal vowels which are phonologically and phonetically different from any SA vowel, which would make the phonetic differences between them more noticeable and easier to detect by SA participants.

5.5.3 English Category Goodness (CG) Contrasts

As for the three English CG contrasts, Table 5-4 shows that two of these were confirmed to be CG contrasts by the results of the present study. First, the English (KIT, DRESS) vowels were analysed as a PAM CG assimilation type because both vowels were found to map to the SA [e] phonetic category with DRESS being closer to that category in terms of categorisations and ratings. Additionally, some of the English KIT productions were mapped to the SA /i/ phonological vowel. Therefore, the PAT of the present study examined the categorisations of the English (KIT, DRESS) vowels into their predicted SA equivalents (/i/, [e]). The results showed that the categorisations of the two English vowels were distributed over the SA phonological category /i/ and its phonetic variants [e] and [ə] with relatively low goodness of fit ratings. Apparently, the participants perceived some phonetic similarity between these English vowels and the three SA categories (/i/, [e], and [ə]), and they could not be decisive about which were the closest equivalents to the English vowels. As a result, the English (KIT, DRESS) contrast was analysed as a CG contrast, in which the two vowels mapped to the three vowel categories grouped together as a single SA phonological vowel /i/. In that case, KIT was found to be closer to the SA vowel category /i/ rather than DRESS.

The phonetic differences between the phonological SA vowel /i/ and its phonetic categories [e] and [ə], which were arguably equivalents to English (KIT, DRESS) vowels, are not as great as the differences between SA phonological /aː/ and its allophone [aːˤ]; see chapter 4 for more details. Therefore, it would be harder for the participants to use the phonetic [e] and [ə] categories in their L2 learning, and they are expected to use one SA phonological category /i/ to identify and discriminate the English (KIT, DRESS) vowels. Therefore, the participants would be expected to misidentify one of them as the other, which will be examined further in the identification task in the next chapter.

The results of the present study also supported the findings of the pilot with regard to the English (GOAT, THOUGHT) vowel contrast. Both vowels were mapped to the SA mid long back vowel /oː/, with THOUGHT being closer to that category in terms of categorisation and goodness of fit rating. Therefore, this contrast has been analysed as a PAM CG assimilation type, and as predicted THOUGHT was categorised as SA /oː/ (68%) significantly more than GOAT (34%). The present study showed that there was a tendency to categorise English GOAT as the SA back vowels (/oː/, /uː/, and /u/) with relatively low goodness of fit ratings, which means that it shares some phonetic properties with all these vowels, i.e. backness.

In the previous chapter, the schwa was analysed in SA as a phonetic vowel category [ə] rather than a phonological one. As a result, SA participants are predicted to have some difficulty perceiving the schwa in English GOAT, which constitutes the first component of this diphthong (/əʊ/). Besides, SA participants might be expected to focus on the second part of this diphthong /ʊ/, which can be categorised to one of their L1/SA phonological back vowels. This argument also applies to the English SQUARE diphthong, in the previous section, which contains a schwa as its second component /ɛə/, which suggests that SA participants would focus on the first component of this diphthong and map it to its closest equivalent in SA. So, this argument provides an explanation for the observed categorisation of English SQUARE as the SA phonological vowel /eː/, and the distribution of English GOAT over the three phonological SA back vowels (/oː/, /uː/, and /u/), with a preference towards the mid long vowel /oː/. The results for the categorisation of English GOAT also support the argument of mapping English diphthongs into SA mid long vowels as a result of extending the phonological assimilation process of vowel-glide sequences of dialectal Arabic into English diphthongs. PAM prediction for the CG English (KIT, DRESS) and (GOAT, THOUGHT) contrasts was that the members of these contrasts will be well discriminated by L2 learners but not as well as the members of the TC assimilation type which will be examined in the following chapter.

5.5.4 English Uncategorised-Categorised (UC) Contrast

The English (FLEECE, NEAR) vowel contrast was predicted to map to the SA high front vowel /iː/, in which FLEECE was analysed as phonologically and phonetically equivalent to that vowel and NEAR was analysed as phonologically close to that category but phonetically uncategorised. Therefore, this contrast was analysed as PAM UC assimilation type, with FLEECE being the categorised vowel and NEAR the uncategorised one. As shown in Table 5-4, the results of the present study supported the findings of the pilot. As predicted, a significant difference was found in the categorisation of the two English vowels, where English FLEECE was categorised as SA /iː/ significantly more often than NEAR was.

Similarly to English SQUARE and GOAT, the uncategorised English vowel NEAR is a diphthong that contains a schwa as its second component. Since schwa was analysed as a phonetic category in SA, the participants' focus was predicted to be on the first part of the diphthong /ɪ/, which is close to a phonological SA category /i/. However, diphthongs were generally associated with SA long vowels. Thus, English NEAR was categorised as a long SA vowel with a high front quality /iː/, which accounts for its high percentage of categorisation (62%) as this particular vowel and for giving it a relatively high goodness of fit rating (5/7). According to the PAM predictions, L2 learners are predicted to have little difficulty discriminating English UC vowels since they are, arguably, able to perceive the phonetic differences between these two vowels.

5.6 Summary

To summarise, the PAT of the present study showed that the SA participants perceived English (TRAP, BATH) and (LOT, STRUT) vowel contrasts as of the PAM TC assimilation type. These contrasts were found to map to SA phonological vs. phonetic vowel categories (/aː/, [aːˤ]) and ([o], /a/). Thus, it was argued that the participants were using their L1 full vowel system including its phonological and phonetic categories.

PAM predicts that the members of these contrasts will have excellent identification and discrimination similar to the ones found in the pilot study.

Also, the English (FACE, SQUARE) contrast was perceived as a PAM SC assimilation type. PAM predicts that the members of this contrast would have poor identification and discrimination since they are mapped to a single SA vowel /eː/. However, the participants were argued to be able to perceive adequate phonetic distance between the English diphthongs, and thus they were able to identify and discriminate these vowels in the pilot study. Therefore, English (FACE, SQUARE) is also predicted to be well identified and discriminated by SA participants in the identification task of the following chapter.

English (KIT, DRESS) and (GOAT, THOUGHT) are analysed as PAM CG contrasts since the members of both of these contrasts were mapped into the SA vowels /i/ and /oː/, respectively, with KIT and THOUGHT being closer to these categories in the number of responses and goodness of fit ratings. PAM predicts a very good identification of the closer English vowels, and a good identification of their counterparts. It also predicts a very good discrimination between the members of these contrasts but not as good as the TC ones.

As for the English (FLEECE, NEAR) contrast, it was perceived as a PAM UC assimilation type as both vowels were perceived as the single SA vowel /iː/, with FLEECE being phonologically and phonetically equivalent to that vowel (in terms of categorisation and goodness of fit ratings), and NEAR being phonologically close to the same SA vowel but phonetically deviant as it received lower goodness of fit ratings. Similarly to the previous contrasts, PAM predicts a very good identification for the categorised vowel FLEECE and a very good discrimination from its counterpart NEAR.

In sum, the present chapter aimed to answer the research questions suggested in section 5.

RQ1: How do SA participants perceptually assimilate English vowels to the five vowel qualities of their SA/L1 phonological inventory?

In general, the results of the present study provided evidence that SA participants used their L1 phonetic allophonic categories in their categorisations of English vowels. For example, English BATH, DRESS, and LOT were categorised as SA phonetic categories [aːˤ], [e], and [o], respectively. However, those phonetic categories did not play a great role in the PAT compared to their phonological counterparts. For example, the schwa was analysed as a phonetic category that does not have any phonological role in SA. As a result, SA participants ignored or did not perceive the schwa in English SQUARE, GOAT, and NEAR diphthongs which have a schwa as one of their diphthongal components. Instead, the participants focused on the other component of each diphthong and used that in their categorisations. On the other hand, the phonetic categories of the L1 are predicted to play a positive role in the identification and discrimination of English vowels which were mapped to in the present experiment. However, this positive role is subject to the phonological relationship between the two SA vowel categories and the phonetic distance between these vowels.

- How do SA participants perceptually assimilate the English tense vs. lax vowels to the long vs. short vowels of SA?

The present study showed that English tense vowels were categorised into SA long vowels. However, a close examination of the categorisation of the above lax English vowels reveals that SA participants showed a tendency to categorise English lax to SA short vowels as well as to SA long ones. For example, the English DRESS, LOT, and STRUT vowels were categorised into L1 long vowels /eː/, /oː/, and /aː/, respectively. Therefore, the categorisation of English monophthongs was analysed in terms of quantity. English monophthongs were grouped into tense vs. lax vowels, and the categorisation of these vowels as long vs. short SA vowels was examined as in Table 5-5 below. A Fisher's exact test was used to examine the distribution of English tense vs. lax vowels into long vs. short L1 vowels. Although a greater number of the lax vowels were categorised as long L1 vowels, the results showed that there was a significant difference in the distribution of the English vowels $p = 0.0001$. Thus, as we would expect, tense vowels are associated mostly with long vowels and lax with short vowels. However, the present study suggests that it is not just a simple mapping of tense-long and lax-short vowels. Other factors might have a role in this categorisation such as vowel intrinsic qualitative properties or language specific characteristics.

Table 5-5: percentages of SA quantitative responses of the classification of SSBE vowels

	SSBE vowels	SA Responses	
		Long vowels /iː/, /eː/, /aː/, /aːˤ/, /oː/, /uː/	Short vowels /i/, /e/, /a/, /ə/, /o/, /u/
Tense vowels	FLEECE, NURSE, BATH, TRAP, THOUGHT, GOOSE	80	20
Lax vowels	KIT, DRESS, STRUT, LOT, FOOT	37	63

- How do SA participants perceptually assimilate English monophthongs vs. diphthongs to the vowels of SA?

The present study revealed that there was a tendency to map English diphthongs, which are phonologically and phonetically different from any SA vowel, into SA mid long vowels /eː/ and /oː/. This mapping was interpreted as an extension of the phonological monophthongisation process of Arabic vowel-glide sequences into mid long vowels. Thus, in their perception of English vowels, SA participants were using not only their L1 phonological and phonetic vowel system but also a historical phonological process from their L1. This might have great implications for perceptual studies of second/foreign language segments. As shown in this study, the effect of the first language extends beyond the effect of its segments and how similar or different they are from the second language segments.

RQ2: To what extent do PAM perceptual categories account for L2 participants' perceptual assimilation of L2 vowels to their L1 vowel inventory?

- How does SA participants' previous phonological knowledge affect their perception of English vowels?

According to the results of the PAT of the present study, the PAM categories explained the categorisation tendencies for part of the data such as the UC contrast (FLEECE, NEAR). However, the previous phonological knowledge of the participants had a great role on their categorisation patterns; and taking into consideration this knowledge had a great effect on the interpretation of the results. Examples of this previous knowledge are the L1 phonetic categories and the monophthongisation phonological process of Arabic.

In the following chapter, the PAM predictions for the identification and discrimination of the above English vowel contrasts are examined in data collected with the same SA participants, compared to a control group of Native English (NE) participants.

6 The Identification Task

The present chapter examines the identification and discrimination of the six English vowel contrasts (TRAP, BATH), (FACE, SQUARE), (KIT, DRESS), (LOT, STRUT), (GOAT, THOUGHT) and (NEAR, FLEECE) which were examined in the PAT of the previous chapter. Unlike the identification task of the pilot study, the task of this study was designed to examine the identification and discrimination patterns of these vowels, in particular. The following research questions were explored again in the light of the results in Chapter 5:

RQ1: How do SA participants perceive the English vowels, in terms of the L2 phonemic vowel inventory, compared to NE participants?

RQ2: Which of these English vowel contrasts are difficult to discriminate by the SA participants?

RQ3: Is this difficulty related to quality or quantity overlap?

RQ4: To what extent can the predictions of PAM assimilation categories account for the results of the present study?

The purpose of the identification study was to test the participants' perception of the above English vowels in terms of the English phonemic vowel inventory, which is used by native English speakers. In particular, the present experiment examines the predictions of the PAM perceptual assimilation categories which were assigned to these English contrasts in the previous experiment. The identification task can help in identifying which English vowel contrasts are difficult to discriminate for SA participants, as compared to NE participants (Tsukada, et al. 2005). It can also help in identifying the source of any discrimination difficulty i.e. temporal or spectral (Ingram and Park 1997).

This chapter is structured as follows: first, the hypotheses for the identification and discrimination patterns of the six English contrasts based on PAM predictions are presented in section 6.1. Then, the methods followed in the identification task are presented in section 6.2, in which any changes from the pilot study identification task will be highlighted. Data analysis procedures are presented in section 6.3. After that, the results of the identification task are presented for each English vowel contrast in section 6.4. In general, as will be seen, the SA participants found it easier to identify and discriminate the members of TC and SC contrasts than CG and UC contrasts. The results of this study are then discussed in terms of PAM categories in section 6.5. Finally, a summary of the analysis of the results of this study is presented in section 6.6, in which the implications of the results of the perceptual experiments of the PAT and the present identification task for the production experiment of the following chapter are indicated.

6.1 Hypotheses

The focus of the identification task was to examine SA participants' identification and discrimination patterns of the six English vowel contrasts under investigation, compared to those of NE participants. The hypotheses for the identification and discrimination patterns of these contrasts were based on the findings of the pilot study and PAM predictions which resulted from the assimilation patterns of these vowel contrasts in the previous chapter. Therefore, the following hypotheses were examined in the present study:

- The members of the English (TRAP, BATH) vowel contrast were categorised as SA phonological /aː/ vs. phonetic category [aːˤ] which are phonologically related. In Chapter 5, the percentage of responses of both English vowels was significantly higher for the SA phonological vowel /aː/ than [aːˤ]. Nevertheless, both English vowels received high goodness of fit ratings with their predicted SA counterparts. Thus, this English contrast was analysed as PAM TC assimilation type. PAM predicts that the members of this contrast will have excellent identification and discrimination by SA participants, which is similar to that of NE participants.
- The members of the second TC English vowel contrast (LOT, STRUT) were also categorised as SA phonetic [o] vs. phonological category /a/, which are phonologically unrelated. Therefore, a significant difference was found in the categorisation of both English vowels, where LOT was mapped into SA [o] and STRUT into SA /a/. Similar to the previous contrast, PAM predicts excellent identification and discrimination for this contrast by SA participants.

- The members of the diphthongal English (FACE, SQUARE) contrast were categorised as a single SA phonological category /eː/, with relatively high goodness of fit ratings. Thus, they were analysed as PAM SC contrast. It was predicted that SA participants applied the historical vowel-glide phonological assimilation process when they perceived these English diphthongs. PAM predicts that the members of the SC contrast will have poor discrimination. However, PAM also suggests that L2 learners do not completely lose their sensitivity to phonetic detail of an L2 contrast that is not similar to any L1 category. This applies to English (FACE, SQUARE) since they are diphthongs and unlike their monophthongal L1 equivalent /eː/ vowel. Therefore, we predict that the participants will be able to perceive the phonetic difference between the two diphthongs, which is suggested to be adequate to have excellent identification and discrimination comparable to that of NE participants.
- As for the English (KIT, DRESS) vowels, the results of the PAT showed that the categorisation of these vowels were distributed over SA phonological category /i/ and its phonetic variants [e] and [ə], with low goodness of fit ratings. Since the participants' categorisations were not decisive, they were grouped together as one SA phonological category /i/. In this case, this English contrast was analysed as the PAM CG assimilation type with KIT being phonologically and phonetically closer to SA /i/ than DRESS. PAM predicts that this contrast will have from moderate to good discrimination, and a very good identification of the closer category KIT.
- Similar to the previous contrast, the English (GOAT, THOUGHT) contrast was also analysed as PAM CG assimilation type. The two vowels were mostly categorised as SA phonological category /oː/, with THOUGHT being closer to that category in the percentage of responses and ratings. Thus, English THOUGHT was considered phonologically and phonetically equivalent to the SA vowel, whereas GOAT was phonetically deviant and thus associated with SA back vowels. PAM also predicts from moderate to good discrimination for this contrast, and a very good identification of the closer category THOUGHT.

- Finally, the English (FLEECE, NEAR) contrast was assimilated as the SA /iː/ phonological vowel category. English FLEECE received a much higher percentage of categorisation than English NEAR. However, both were given high goodness of fit ratings. As a result, this contrast was analysed as the PAM UC assimilation type, in which FLEECE was the categorised vowel as it was phonologically and phonetically equivalent to the SA /iː/ vowel, whereas English NEAR was uncategorised as it was phonologically equivalent to the SA category but phonetically deviant. The members of this contrast are predicted to receive very good identification and discrimination, with better identification for the categorised English vowel FLEECE.
- In sum, the identification and discrimination patterns of the above English contrasts in terms of PAM predictions will be in this order: TC ≥ SC > UC > CG, i.e. from excellent to moderate.

6.2 Methodology

The productions from the same English male speaker used in the pilot study and in the PAT of the previous chapter were used in the identification task, following the same recording procedures and methods. For more information about the English speaker and the recording procedures see Section (3.1.5).

The same ten female and ten male Syrian Arabic participants who took part in the PAT in the previous chapter also took part in the identification experiment; for more information about the SA participants see Section 5.2. As a control group, five English female and four male speakers participated in the identification task. Similarly to the SA participants, a questionnaire was used to obtain information about the English participants and their language background. The English participants reported themselves as speakers of Standard Southern British English (SSBE). They were chosen from students and staff at the University of York. Their average age was 32.8 years old. The English participants were paid for their participation.

Unlike the identification experiment in the pilot study, which was designed to explore the identification and discrimination patterns of all English vowels, the identification task in the present study was designed to examine only the identification and discrimination of the six English vowel contrasts presented in Table 6-1 below. The table shows that, similarly to the pilot, the stimulus material for this experiment consisted of a list of real monosyllabic /CVC/ words. All target words were presented in the same phrase used in the PAT 'say ____ again'. Similarly to the previous experiments, the English speaker produced three blocks of stimuli. Following the same procedures, the best productions were extracted and judged by the same two English judges (see Section 3.1.4).

Table 6-1: English vocalic contrasts which were presented to SA and NE participants to obtain identification results. The shaded cells show the actual target words which were presented to the participants.

same responses

English contrast	Option 1	Option 2	Option 3	Option 4	Option 5	presented
(TRAP, BATH)	TRAP	BATH	GOAT	STRUT	LOT	
	Cat	card	coat	cut	cot	10 times
(LOT, STRUT)	LOT	STRUT	GOAT	THOUGHT		
	cot	cut	coat	caught		8 times
(FACE, SQUARE)	FACE	SQUARE	NEAR	FLEECE	DRESS	
	phase	fares	fears	fees	fez	10 times
(KIT, DRESS)	KIT	DRESS	FLEECE	NEAR		
	bit	bet	beat	beard		8 times
	DRESS	KIT	FLEECE	FACE		
	bet	bit	beat	bait		8 times
(GOAT, THOUGHT)	GOAT	THOUGHT	STRUT	LOT		
	coat	caught	cut	cot		
(FLEECE, NEAR)	FLEECE	NEAR	FACE	KIT	DRESS	
	fees	fears	phase	fizz	fez	10 times
Total						**54**

Table 6-1 presents the vocalic contrasts which were used as responses in the first part of the identification task. The actual target words in the shaded cells were chosen as minimal pairs which differed in their vowel only, in most cases. The list was chosen to consist of simple and familiar words to SA participants to avoid misidentification for reasons other than vowel category[8]. Each set of responses were repeated eight or ten times (depending on the number of responses x 2), with each of the responses presented twice as the target choice. These contrasts were randomised, within each set, and presented to the participants. This part of the task aimed to obtain identification results when there are distracters other than the target words, e.g. in Table 6-1, the target words in the first set of responses for the (TRAP, BATH) contrast are (cat and card), whereas the other words (coat, cut, and cot) were the distracters. These distracters were chosen to share some characteristics with the target contrasts such as place, height, or roundness.

The identification task was immediately followed by paired presentations of the words representing the target English vowel contrasts, e.g. cat and card for (TRAP, BATH), paired with each other and with the other distracters from Table 6-1. Table 6-2 below shows an example of the paired presentation of (TRAP, BATH) set, this set was repeated four times with each of the responses heard twice. Paired presentation aimed at narrowing the focus on the identification and discrimination of the vowel contrasts under investigation, without any distracters, to obtain discrimination results. All stimuli were presented within two blocks: one for the basic contrasts with four or five responses (Table 6-1), and the other for the paired contrasts with only two responses (Table 6-2). All together, 162 items were presented to the participants.

Table 6-2: An example of the English paired vocalic contrasts which were presented to SA and NE participants to obtain discrimination results of (TRAP, BATH) contrast.

TRAP	BATH
(cat, card)	
(cat, coat)	(card, coat)
(cat, cut)	(card, cut)
(cat, cot)	(card, cot)

For the Syrian Arabic participants, the identification experiment was run in a quiet computer room at the Asia Institute for Languages in Damascus. Similar to the pilot, the experiment was run using the Praat MFC experiment (Boersma and Weenink 2009). The

[8] Word familiarity was judged based on my knowledge of English educational materials in Syria, where I taught English for three years at school and university levels.

headphones were Sennheiser HD 280 Pro 64 Ω. Spelling, pictures and pronunciation were used to familiarise the participants with the English words.

After that, the participants were asked to choose one of five, four or two alternatives presented on flash cards in a forced-choice identification task. These flash cards were used instead of presenting the alternative responses on a response sheet as in the pilot study. Each card contained the pictures of the alternative words with their written form in conventional English orthography. This was a self-paced task; the listeners were directed to listen to the phrase "say ___ again", choose the word from the flash card that best matches the target word, and then press the (A, B, C, D, or E) alternative on the computer screen (as can be seen from the screen shot of the experiment in Figure 6-1 below). The participants had the five-alternative response screen, but they were directed to focus on the first two if they had paired responses. As in the pilot, a replay button was added so the participants can listen to the words as many times as they wished. Finally, when they had decided on the target word they pressed the OK button and moved to the next item.

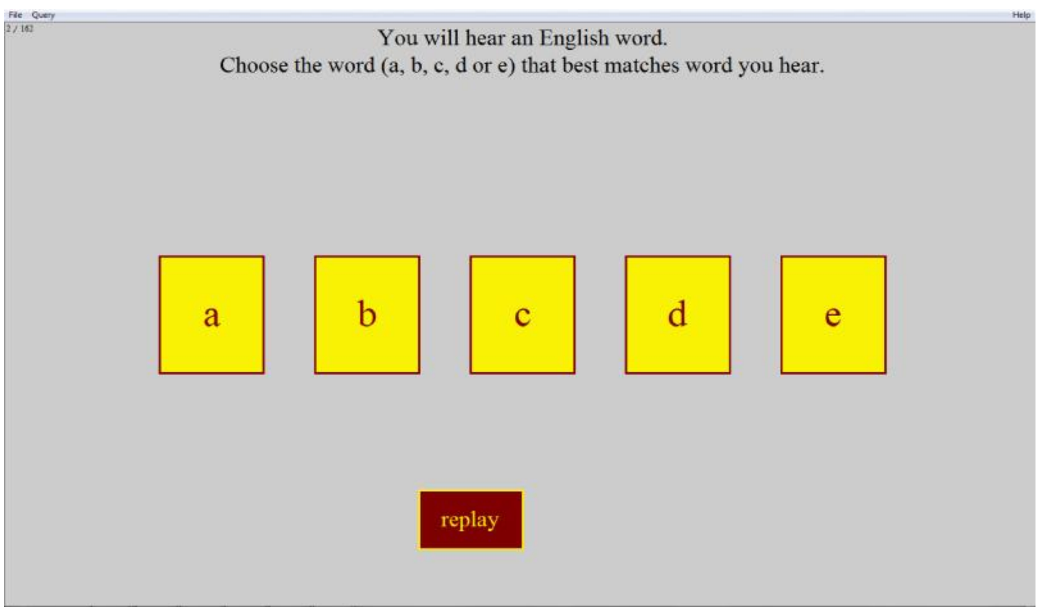

Figure 6-1: The identification experiment as presented to the participants using Praat MFC experiment, this was presented simultaneously with flash cards for SA participants, and with a response sheet for the NE participants.

For the English participants, the experiment was run in the recording room in the department of Language and Linguistic Science at the University of York. The experiment was run following the same procedures used for the SA participants. However, using pictures for the alternative words was not necessary for NE participants as it was for SA participants. Thus, the alternatives were presented on a response sheet using English orthography, and the participants were asked to choose the target word from the sheet instead of the flash cards.

6.3 Data Analysis

A matrix of errors of identification similar to the one used in the pilot study was created for each participant. However, the focus of the identification of the present study was to explore the identification of the six English vowel contrasts which were analysed in the PAT of the previous chapter. Table 6-3 below presents the percentages of the correct and incorrect identification for these English vowels across all SA participants. The results of the identification and discrimination parts of the study are merged in this table and presented as percentages of identification. The shaded cells presents the discrimination results of the target vowels. A similar table was created for NE participants, whose identifications were 100% correct for the same vowels. Therefore, only the results of the SA participants are presented here.

Table 6-3: Percentages of correct (shaded cells) and incorrect identification of the six English vowel contrasts by SA participants.

	TRAP	BATH	LOT	STRUT	FACE	SQUARE	KIT	DRESS	GOAT	THOUGHT	FLEECE	NEAR	**NUM**
TRAP	94	1		5									**200**
BATH	4	90	2	2					2				**200**
LOT	1	1	51	4					27	16			**360**
STRUT	2		3	95									**280**
FACE					95		2				3		**240**
SQUARE						98		2					**200**
KIT					2		68	26			2	2	**320**
DRESS					3		6	89			2		**280**
GOAT			13						73	14			**440**
THOUGHT									23	77			**200**
FLEECE							2	1			94	3	**280**
NEAR					4		7	5			11	73	**280**

Three two-sided Fisher's exact tests were conducted for each English vowel contrast. The first test was run to examine the differences between the SA and NE participants in terms of the number of correct identifications between the members of each English contrast, as shown for (TRAP, BATH) contrast in Table 6-4 below. The second test was run to examine the difference between the members of each contrast with regard to the number of correct vs. incorrect identifications. Finally, the discrimination of the members of these contrasts was also examined in a third test. The difference between the number of correct identifications of TRAP and BATH vowels vs. misidentifications of each vowel as the other was calculated for SA speakers. Following the previous chapter, the Fisher's exact test was used because it is suitable for the categorical data of the present study, which might have less than five responses in some of its cells as shown in Table 6-3 above. The results of these tests are presented and discussed for each English vowel contrasts in the following sections.

Table 6-4: Representation of the Fisher's exact tests for English (TRAP, BATH) contrasts. (1) examines the difference between SA and NE participants in the number of correct identifications of English (TRAP, BATH) vowels.

Fisher's exact test 1		
	TRAP	BATH
SA	correct Id	correct Id
NE	correct Id	correct Id

6.4 Results

6.4.1 The Identification of English (TRAP, BATH) contrast

Figure 6-2 below presents the percentage of English (TRAP, BATH) identification by SA female and male participants. As can be seen, female and male participants identified the English vowels very accurately and there are no differences between them. Figure 6-3 presents the same results averaged across all the SA participants. It shows that both English vowels received high correct identification close to that of NE participants, who had 100% correct identification for both vowels. A Fisher's exact test was used to examine the difference between SA and NE participants in the number of correct identifications of English (TRAP, BATH) vowels. As expected, the result showed that there was no significant difference between SA and NE participants in the correct identification of this contrast: $p = 0.828$.

Another Fisher's exact test was run to examine the difference in the number of correct vs. incorrect identification of English (TRAP, BATH) vowels by the SA participants. As shown in Figure 6-3, the results showed that there was no significant difference in the identification of the two vowels: $p = 0.275$. Also, SA participants' discrimination of these vowels was examined in a third Fisher's exact test. In this test, the difference between TRAP and BATH vowels was calculated in terms of their correct identification vs. Misidentification one as the other. As predicted, there was a significant difference in the correct identification vs. misidentification (TRAP as BATH and vice versa) of these English vowels: $p = 0.0001$, which means that English TRAP was mostly correctly identified as TRAP, and English BATH was mostly correctly identified as BATH. Thus, the two English vowels had excellent discrimination by SA participants close to NE.

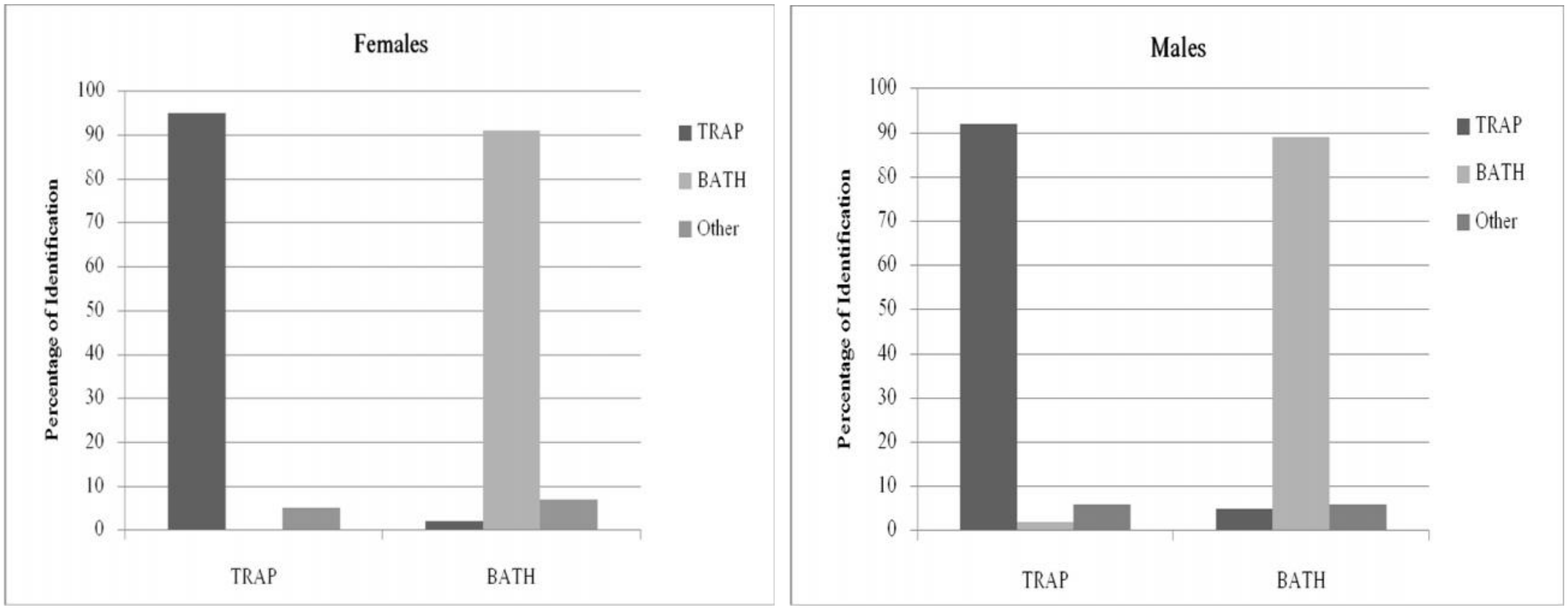

Figure 6-2: SA females and males' percentages of identification of English (TRAP, BATH) vowels as English TRAP, BATH, or other vowel categories grouped together.

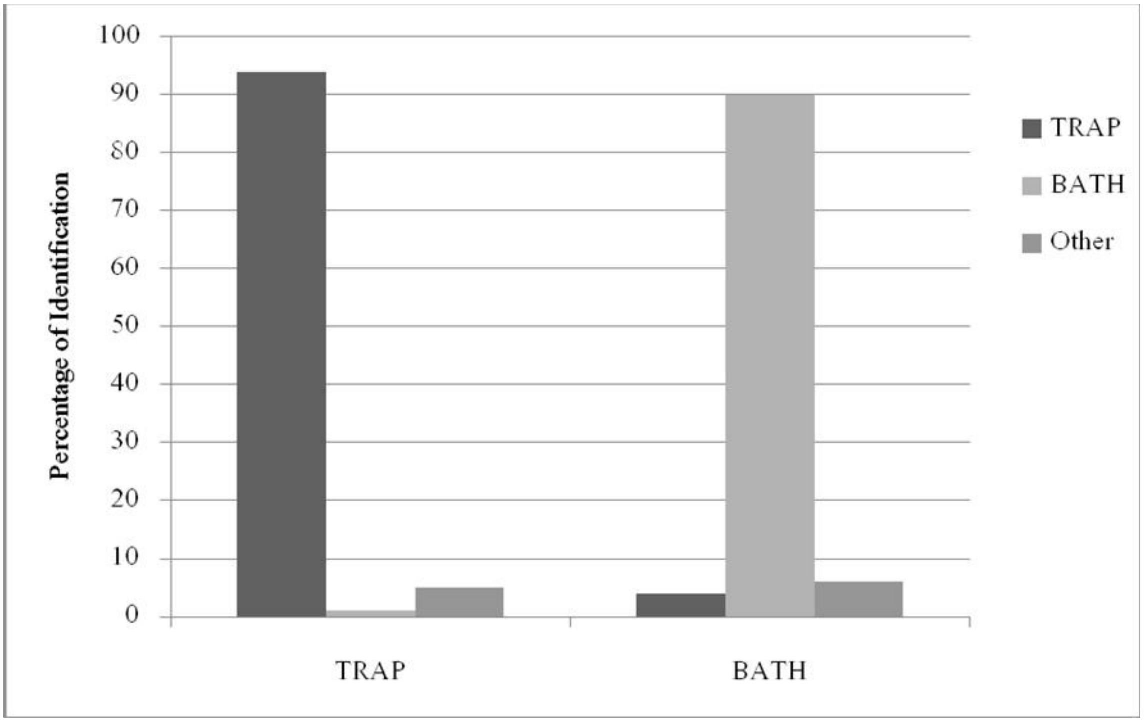

Figure 6-3: Overall percentage of identification of English (TRAP, BATH) vowels as English TRAP, BATH, or other vowel categories grouped together, across all SA participants.

The hypothesis for the identification of the TC English (TRAP, BATH) contrast by SA participants was that the members of this contrast will have excellent identification and discrimination close to NE. This hypothesis was confirmed by the results of this study as the two members of this contrast were very accurately identified and discriminated by SA participants. These results suggest that although SA participants did not directly categorise a great number of English BATH vowels, in the PAT, into their predicted SA equivalent, i.e. the pharyngealised allophone [aːˤ], they may have been able to use their L1 phonetic category to identify a phonetically similar English vowel, which is BATH.

6.4.2 The Identification of English (LOT, STRUT) contrast

Similarly to the previous contrast, the percentages of females and males' identification of the second English TC contrast (LOT, STRUT) are presented in Figure 6-4 below. The identification patterns of females and males look very similar, in which English STRUT is very accurately identified with a percentage close to that of NE participants (who had 100% correct identification). On the other hand, the English LOT vowel received a low percentage of correct identification by both males and females. A Fisher's exact test was run to examine the difference between SA and NE participants in the correct identification of these two vowels. The results showed that NE participants correctly identified the members of this contrast significantly more than SA participants: $p = 0.0001$. A closer examination of Figure 6-5 below shows that the difference lies in the low identification of English LOT vowel by SA participants. Only 51% of LOT was correctly identified, with 49% identified as other English vowels.

The results of the second Fisher's exact test examining the correct vs. incorrect identification of these vowels by SA participants confirmed that the participants indeed had significantly higher correct identification for STRUT than for LOT: $p = 0.0001$. Further examination of the results revealed that 27% of English LOT was misidentified as English GOAT, and 16% as English THOUGHT. As predicted for the TC contrast, the results of the third Fisher's exact test showed there was a significant difference in the correct identification vs. misidentification of the members of this contrast: $p = 0.0001$, which means that the (LOT, STRUT) vowels had excellent discrimination as they were not misidentified as each other.

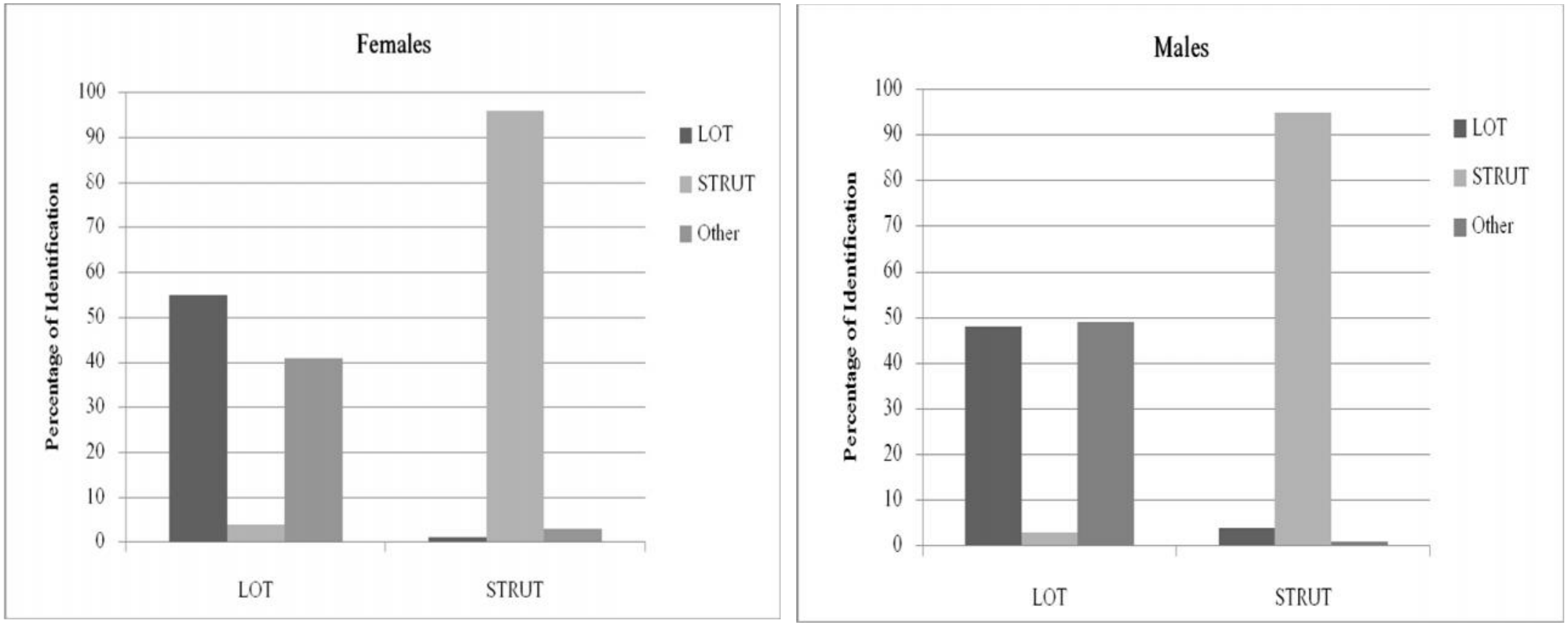

Figure 6-4: SA females and males' percentages of identification of English (LOT, STRUT) vowels as English LOT, STRUT or other vowel categories grouped together.

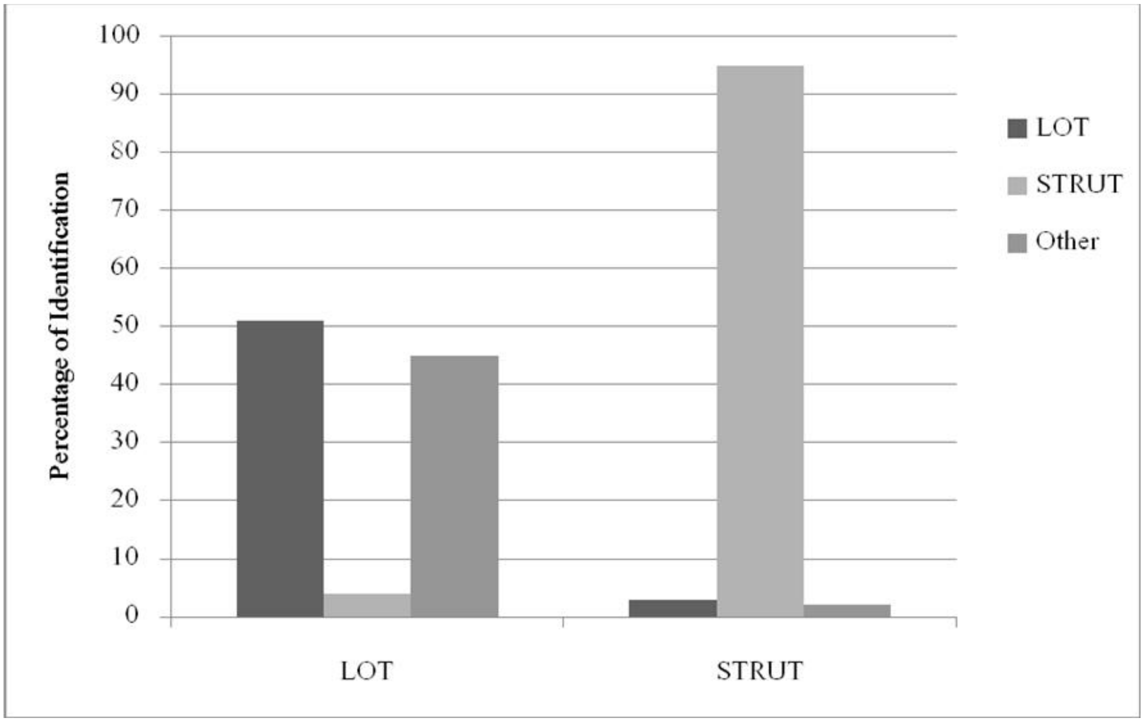

Figure 6-5: Overall percentage of identification of English (LOT, STRUT) vowels as English LOT, STRUT, or other vowel categories grouped together, across all SA participants.

Similarly to the previous contrast, the hypothesis for the identification of the TC English (LOT, STRUT) contrast was that the members of this contrast were predicted to have excellent identification and discrimination since they were mapped into two separate SA vowel categories ([o], /a/). The results of the present study partially support this hypothesis, in which English (LOT, STRUT) had excellent discrimination and excellent identification of the STRUT vowel.

6.4.3 The Identification of the English (FACE, SQUARE) contrast

As for the the English SC (FACE, SQUARE) contrast, Figure 6-6 below shows that SA female and male participants had excellent identification of both vowels, which is close to NE participants': $p = 0.837$. As predicted, Figure 6-7 shows that across all SA participants English (FACE, SQUARE) had similarly high correct identification: $p = 0.082$, as well as excellent discrimination: $p < 0.0001$.

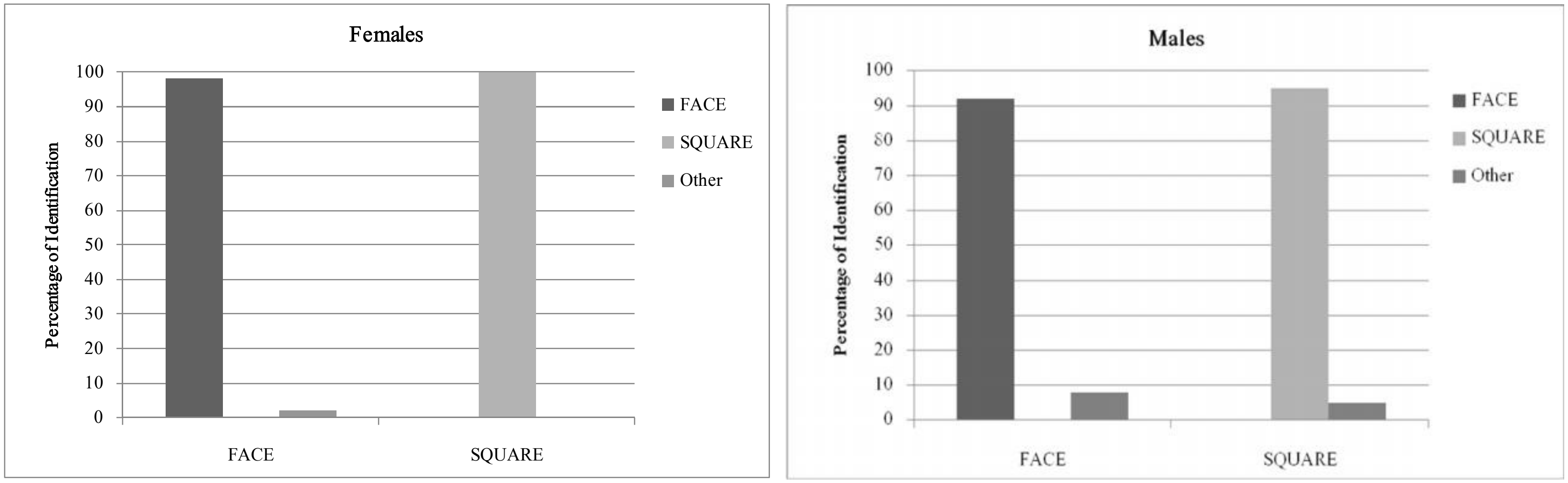

Figure 6-6: SA females and males' percentages of identification of English (FACE, SQUARE) vowels as English FACE, SQUARE, or other vowel categories grouped together.

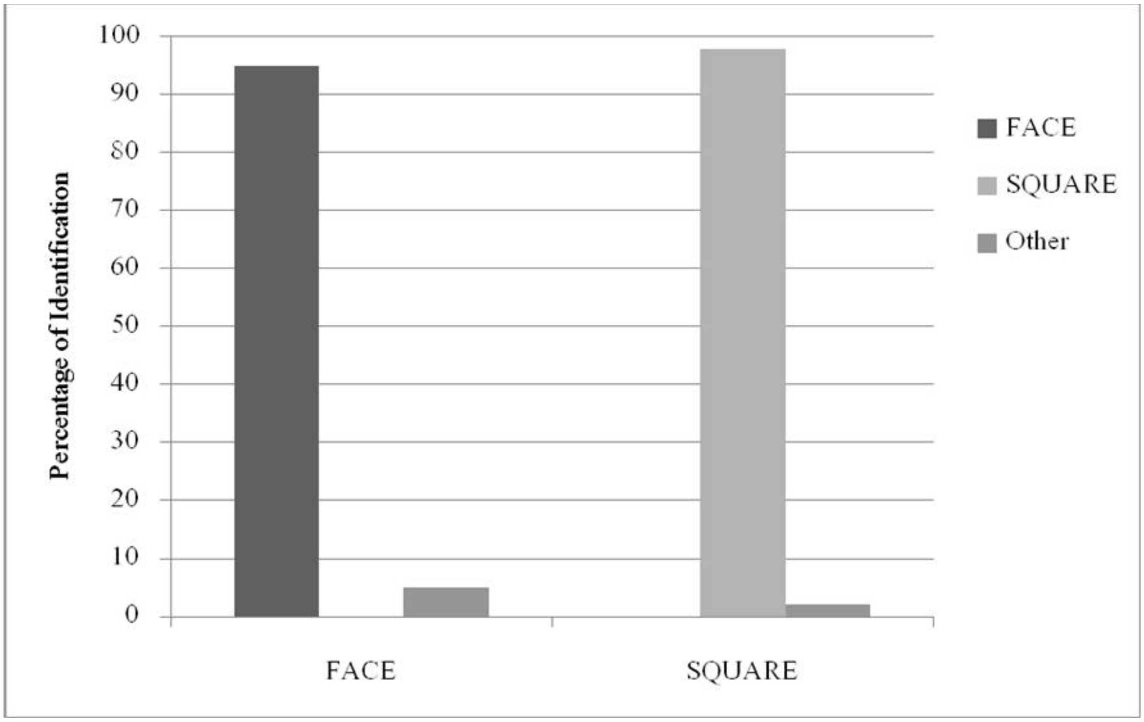

Figure 6-7: Overall percentage of identification of English (FACE, SQUARE) vowels as English FACE, SQUARE, or other vowel categories grouped together, across all SA participants.

PAM originally predicts poor discrimination for the members of SC contrast. However, since PAM suggests that learners do not lose their phonetic sensitivity to L2 contrasts, particularly for such a prevailing phonetic difference between FACE and SQUARE, I hypothesized that the SA participants might be able to perceive enough phonetic differences between these two diphthongs. Thus, English (FACE, SQUARE) vowels were expected to have excellent identification and discrimination similar to NE participants. The results of the present study strongly support this hypothesis since both vowels were very accurately identified and discriminated by SA participants. Also, those identification and discrimination patterns were not significantly different from those of NE participants. These findings support the earlier argument that the SA participants did perceive adequate phonetic distance between these two vowels, which was enough to identify and discriminate them.

6.4.4 The Identification of the English (KIT, DRESS) contrast

Similar to the previous contrasts, Figure 6-8 below shows that SA female and male participants had similar identification patterns for the English CG (KIT, DRESS) contrast. However, Figure 6-9 below shows that SA participants’ correct identification of these vowels were not as high as those of the NE participants (which were 100% correct identification for both vowels), which was supported by the results of the Fisher’s exact test: $p = 0.036$. The difference between SA and NE participants can be attributed mainly to the low percentage of correct identification of English KIT vowel. This was also supported by the Fisher’s exact test, which showed that English DRESS vowel had a significantly higher correct identification than English KIT: $p = 0.0001$. Further examination of the results revealed that 26% of English KIT was misidentified as its counterpart DRESS vowel. However, the results of the Fisher’s exact test, comparing the correct identification of these vowels vs. Misidentifying them as each other, showed that the two vowels had significantly very good discrimination: $p = 0.0001$.

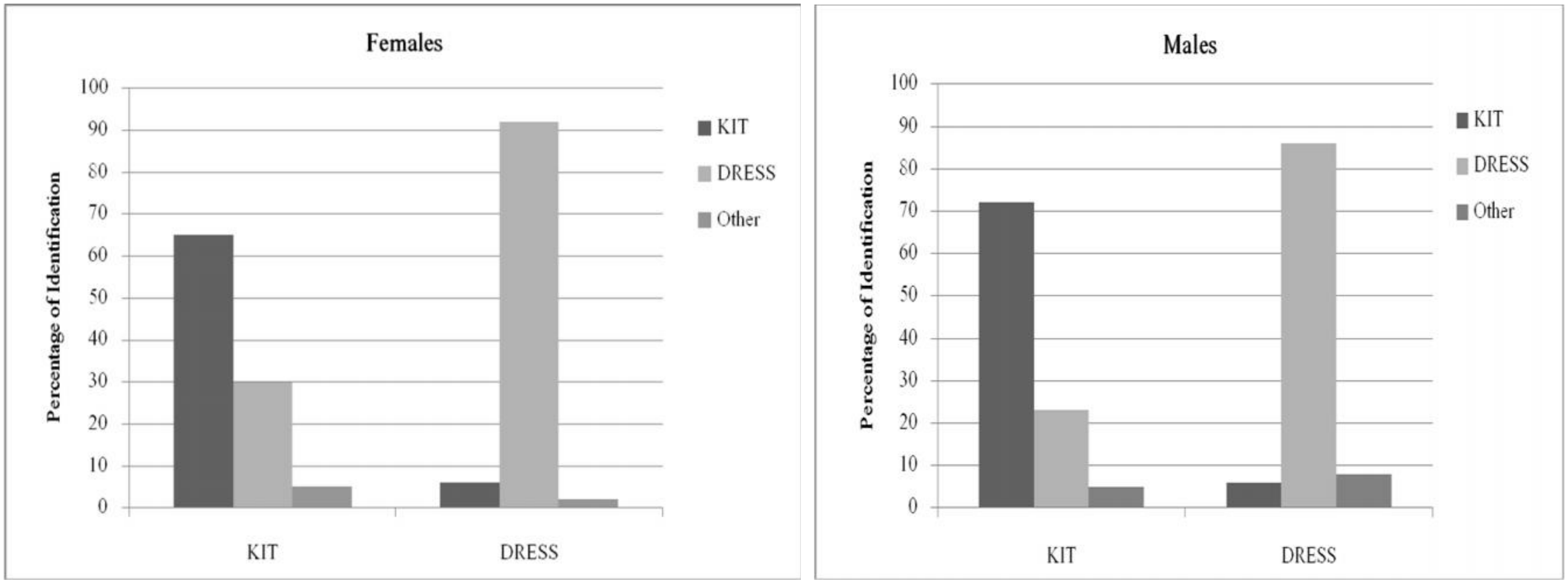

Figure 6-8: SA females and males' percentages of identification of English (KIT, DRESS) vowels as English KIT, DRESS, or other vowel categories grouped together.

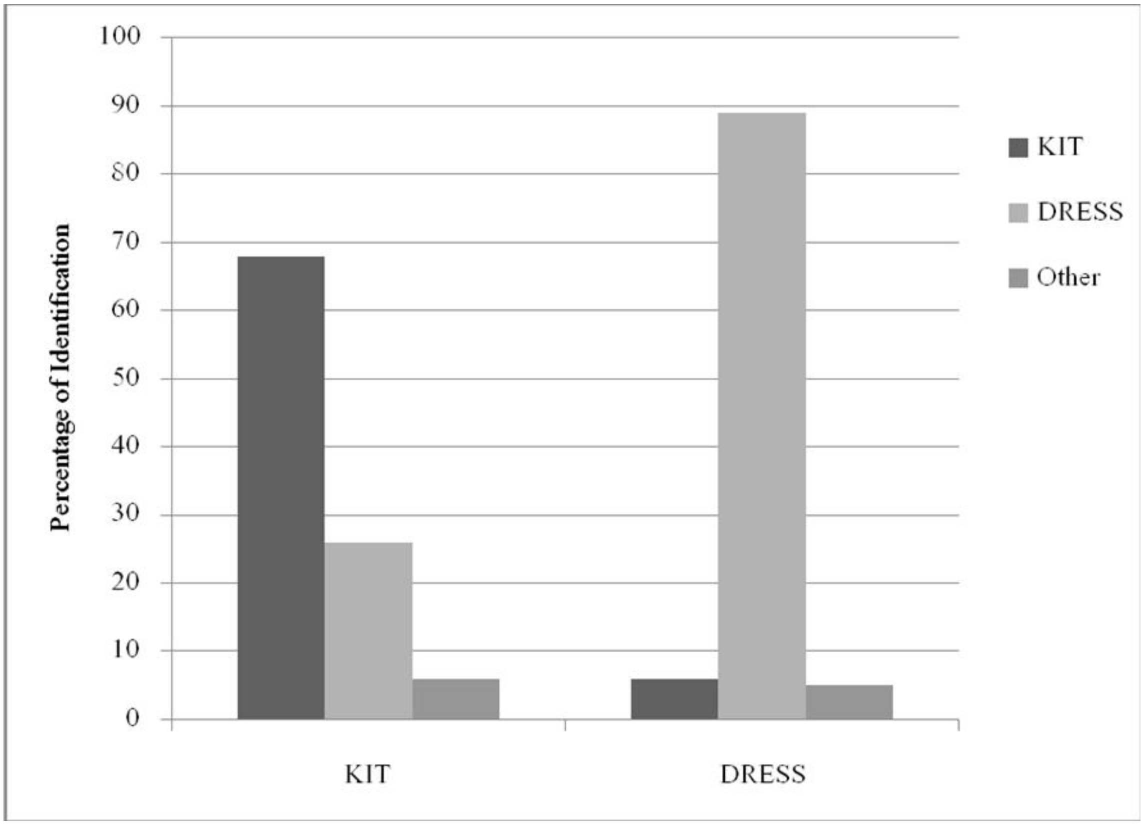

Figure 6-9: Overall percentage of identification of English (KIT, DRESS) vowels as English KIT, DRESS, or other vowel categories grouped together, across all SA participants.

The hypothesis for the identification of the English CG (KIT, DRESS) contrast was that the members of this contrast were predicted to have from moderate to good discrimination since they were mapped into a single SA phonological category /i/. Also, a very good identification was predicted to English KIT as it was found to be closer to the L1 category in the previous chapter. The results of the present study support the first part of the hypothesis since the English (KIT, DRESS) vowels had significantly very good discrimination by SA participants. However, the identification of English KIT was not as good as predicted. Additionally, the identification of English DRESS was significantly better, and close to that of the NE.

6.4.5 The Identification of the English (GOAT, THOUGHT) contrast

Figure 6-10 below shows that SA female and male participants have the same identification patterns for the second English CG (GOAT, THOUGHT) contrast. Across all participants, Figure 6-11 shows that both English vowels have very good identification, but not as good as NE (100% correct identification for both vowels). However, the results of the Fisher's exact test showed that the number of correct identifications of SA participants' were not significantly different from those of the NE for both vowels: $p = 0.745$. Additionally, there was no significant difference in the correct vs. incorrect identification of these vowels by SA participants: $p = 0.379$. Figure 6-11 shows that 14% of English GOAT was misidentified as THOUGHT, and 23% of English THOUGHT was misidentified as GOAT. Nevertheless, there was a significant difference in the identification of these vowels: $p = 0.0001$, which means they were correctly identified and had very good discrimination.

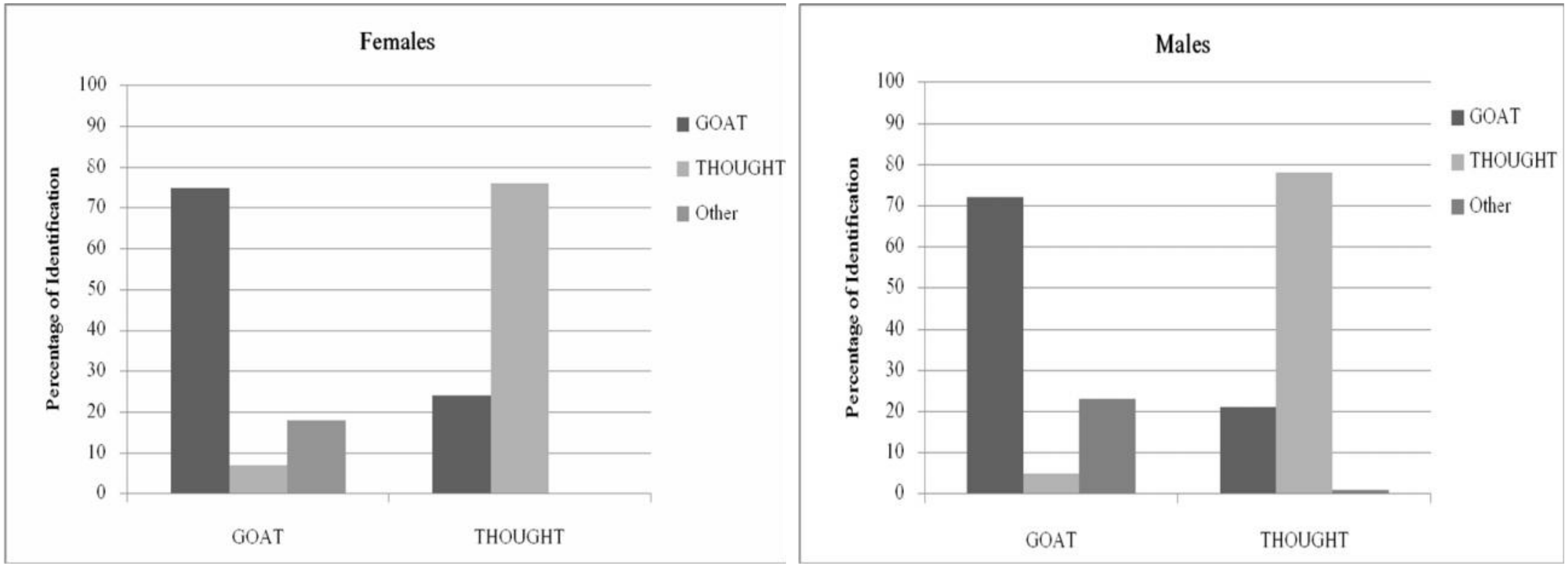

Figure 6-10: SA females and males' percentages of identification of English (GOAT, THOUGHT) vowels as English GOAT, THOUGHT, or other vowel categories grouped together.

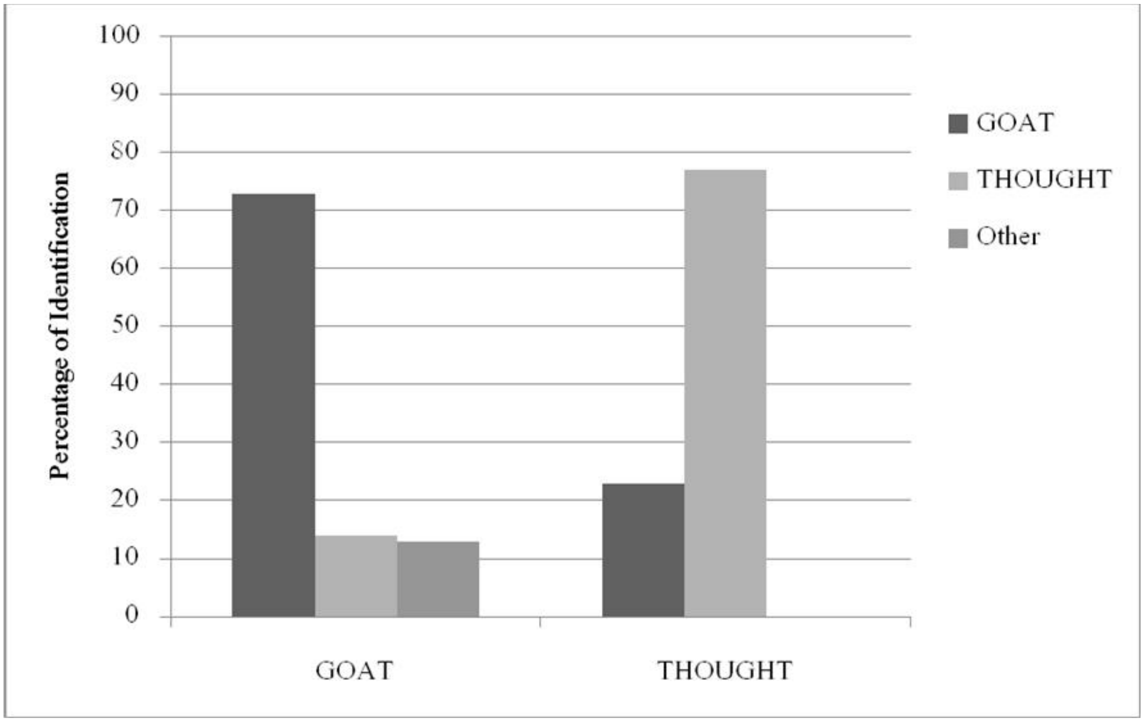

Figure 6-11: Overall percentage of identification of the English (GOAT, THOUGHT) vowels as English GOAT, THOUGHT, or other vowel categories grouped together, across all SA participants.

Similar to the previous CG contrast, the members of the English (GOAT, THOUGHT) contrast were predicted to have from moderate to good discrimination and a very good identification of English THOUGHT vowels, which was found to be closer to the SA equivalent vowel /oː/. The results of the present study confirmed the hypothesis for this contrast. The two members of this contrast were found to have very good discrimination, as well as, a slightly (but not significantly) better identification of English THOUGHT vowel, see Figure 6-11 above. In the previous chapter, the English diphthong GOAT was analysed as phonetically deviant from its equivalent SA phonological category /oː/; therefore, it was mapped into SA back vowels, in general. This analysis was also reflected in the identification of this vowel as it was misidentified 13% as English back vowel LOT, and 14% as English back vowel THOUGHT.

6.4.6 The Identification of the English (FLEECE, NEAR) contrast

Figure 6-12 below shows similar identification patterns for the UC English contrast (FLEECE, NEAR) by both SA female and male participants. However, females' correct identification of English NEAR was less than males', and it was misidentified as English FLEECE more by females than by males. In general, the identification patterns across all SA participants show that English FLEECE was correctly identified significantly more often than English NEAR: $p = 0.0001$, as shown in Figure 6-13 below. As a result, the number of correct identifications of these vowels was at marginal level of significance less than that of NE, $p = 0.059$. However, this difference was attributed to the low correct identification of the English NEAR vowel. As for the discrimination of these vowels, as predicted, they had significantly very good discrimination: $p = 0.0001$ as they were both correctly identified by SA participants.

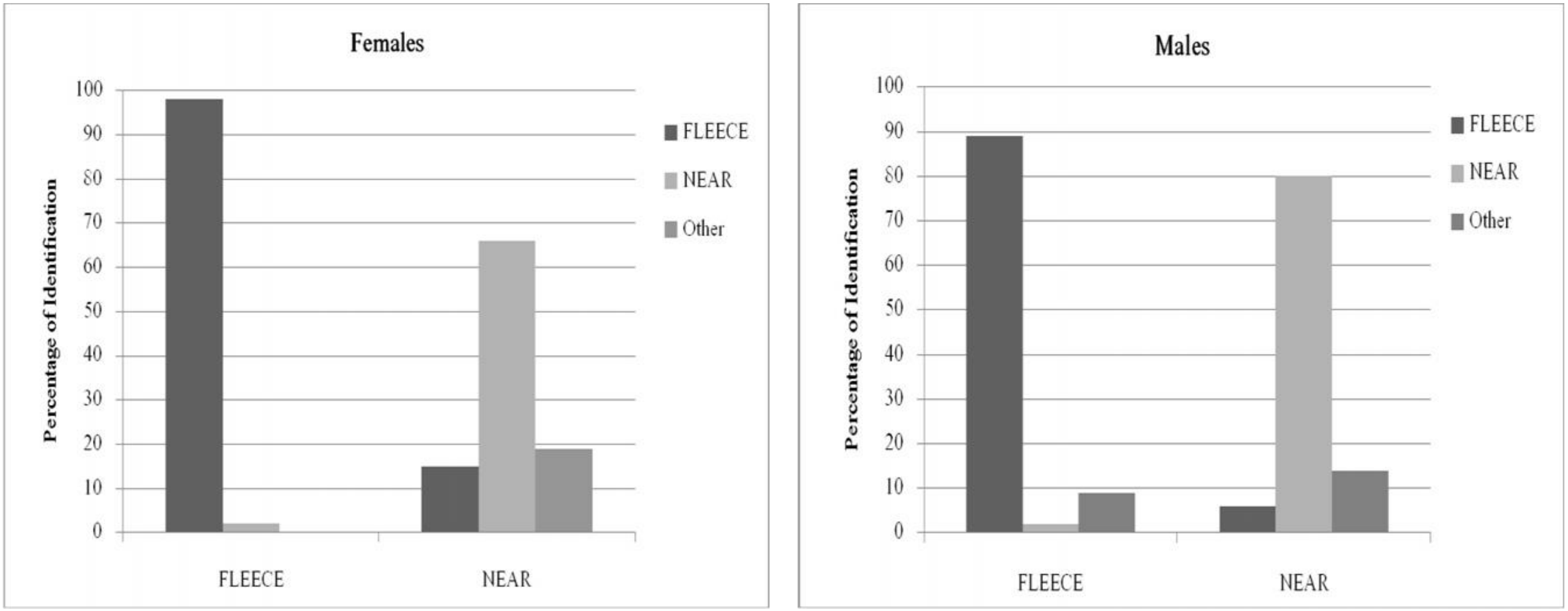

Figure 6-12: SA females and males' percentages of identification of English (FLEECE, NEAR) vowels as English FLEECE, NEAR, or other vowel categories grouped together.

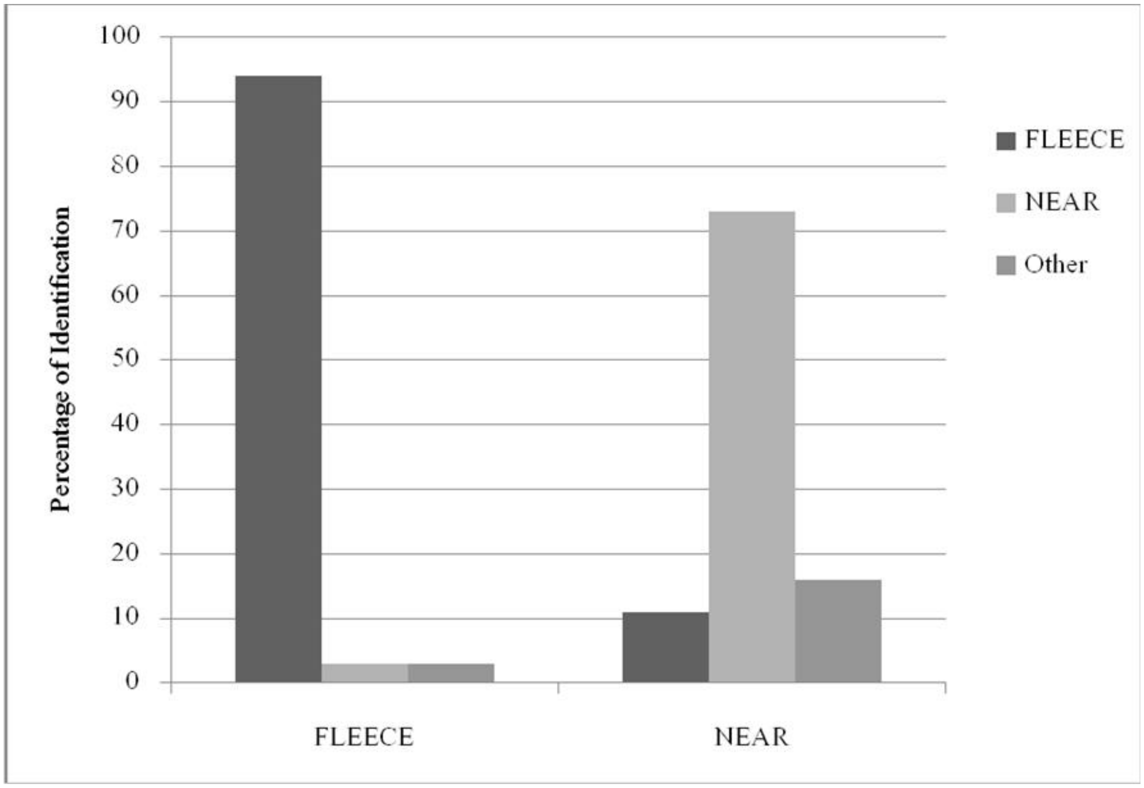

Figure 6-13: Overall percentage of identification of English (FLEECE, NEAR) vowels as English FLEECE, NEAR, or other vowel categories grouped together, across all SA participants.

The hypothesis for this contrast was that the members of the UC contrast were predicted to have very good identification and discrimination, with a better identification for English FLEECE, which was analysed as the categorised vowel in the previous chapter. The results of the present study support this hypothesis since the English (FLEECE, NEAR) vowels had very good discrimination and a better identification of English FLEECE.

6.4.7 The Identification of PAM categories

In sum, the above results for the identification and discrimination of the six English vowel contrasts by SA participants showed that the participants had excellent identification and discrimination of the TC, except for the LOT vowel which had poor identification (51%). On average, the percentage of correct identification for the TC contrasts (TRAP, BATH, LOT, STRUT) is 82.5%. However, without LOT, which had poor identification, the percentage of correct identification is much higher at 93%.

Similarly, the SC contrast showed excellent identification and discrimination which was, on average across both vowels (FACE, SQUARE), 96.5%. The CG contrasts, on the other hand, showed less correct identification and discrimination than the previous PAM categories. On average and across (KIT, DRESS, GOAT, THOUGHT), they showed 76.78% correct identification. Finally, the UC contrast (FLEECE, NEAR) had on average 83.5% correct identification across both vowels.

Best et al. (2001) predicts a gradient perception of TC>CG>SC contrasts. However, based on the assumption that the participants were able to perceive the phonetic difference between the members of the diphthongal SC contrast, I hypothesized that the SC will have a similar percentage of identification or slightly lower than the TC. Additionally, the UC contrast was added to this predicted gradient perception, in which I predicted it will have better identification than the CG contrasts but less than the TC and SC contrasts, i.e. the predicted order of identification is TC ≥ SC > UC > CG from excellent to moderate. The results of the identification task showed that the SC contrast showed a slightly better identification than the TC contrasts (if LOT was excluded), and both showed better identification than the UC contrast, which was better than the CG contrasts, i.e. the percentages were SC (96.5%)> TC (93% (82.5% with LOT)) ≥ UC (83.5%)> CG (76.78%).

6.5 Discussion

In general, the results of the identification task of the present study showed that SA participants did not have great difficulty in identifying or discriminating most of the members of the six English vowel contrasts under investigation. Table 6-5 below presents a summary of these results, which included comparing SA participants' correct identification of these contrasts to that of NE participants. Additionally, the results included examining the identification and discrimination of the members of these contrasts by SA participants.

The English vowel contrasts presented in Table 6-5 were analysed in the previous chapter in terms of PAM perceptual assimilation categories. Based on the results of the PAT of the previous chapter, these vowel contrasts were assigned one of the PAM assimilation types, which yielded specific predictions for the identification and discrimination patterns for the members of those contrasts by second/foreign language learners, i.e. SA participants. Therefore, the identification task examined the identification and discrimination of those English contrasts with specific reference to the PAM predictions, except (FACE, SQUARE) contrast which had its predictions based on the results of the pilot study. In the following sections, the results of this task are discussed within PAM perceptual assimilation categories.

Table 6-5: summary of the SA participants' identification and discrimination results of the English vowel contrasts under investigation compared to those of NE participants

English contrast	PAM type	Compared to NE	Identification	Discrimination
(TRAP, BATH)	TC	similar	(94%, 90%)	excellent
(LOT, STRUT)	TC	different	(51%, 95%)	excellent
(FACE, SQUARE)	SC	similar	(95%, 98%)	excellent
(KIT, DRESS)	CG	different	(68%, 89%)	moderate
(GOAT, THOUGHT)	CG	different	(73%, 77%)	good
(FLEECE, NEAR)	UC	different	(94%, 73%)	very good

6.5.1 English Two Category (TC) contrasts

In the previous chapter, English (TRAP, BATH) and (LOT, STRUT) vowel contrasts were categorised as SA phonological vs. phonetic vowel categories, i.e. (/aː/, [aːˤ]) and ([o], /a/). Therefore, they were analysed as the PAM TC assimilation type, which predicts that the members of these contrasts will have the best identification and discrimination patterns by SA participants, comparable to NE patterns. Table 6-5 above shows that the results of the present study supported PAM predictions for these contrasts, except for the moderate identification of English LOT vowel which was significantly lower than NE identification. However, English LOT was not confused with its counterpart STRUT; therefore, they had excellent discrimination as expected by PAM.

The difficulty in the identification of English LOT can be directly mapped onto the difficulty which the SA participants had in the categorisation of this vowel. In the previous chapter, SA participants distributed the categorisation of English LOT into their L1 mid back vowels /oː/ (32%) and [o] (26%), which suggests that English LOT represents to the SA participants a mid-back vowel quality, which is perceived with longer vowel duration than their short L1 vowel [o]. Thus, this suggestion can account for the participants' misidentification of English LOT as other English mid-back long vowels, i.e. GOAT and THOUGHT.

The categorisation of English BATH was also distributed over two SA categories (/aː/, [aːˤ]), in this case phonologically-related with more responses into the phonological category /aː/. Being categorised into a phonologically related SA contrast, with more responses into the phonological category, which is the canonical form to represent both the phonological and the phonetic categories, should be expected. This is because naive speakers of any language are not expected to be aware of the phonetic and allophonic categories of their language. On the other hand, naive speakers are expected to be aware of their L1 phonemic and phonological categories, which are taught directly in schools. Nevertheless, this canonical representation form was deactivated when the participants were asked to identify and discriminate the English BATH vowel in the present study. In the identification task, the participants were able to subconsciously access their L1 phonetic vowel categories, and use them to identify phonetically equivalent English vowels. This argument suggests that SA participants used their L1 phonetic allophonic categories [aːˤ] and [o] to identify English BATH and LOT vowels, respectively, which are relevant from an articulatory-phonetic perspective.

One can argue that using L1 phonetic allophonic categories in the identification and not in the categorisation can be due to the way we access our phonetic vs. phonemic categories, which might be in the subconscious vs. conscious level of awareness. This means that the participants were using their L1 phonological vowel space when they were asked directly to map English vowels into their closest SA equivalent vowels. However, when the participants were asked to identify and discriminate these English vowels, they have, arguably, accessed their L1 phonological and phonetic vowel space.

6.5.2 English Single Category (SC) contrast

In the PAT, the two English diphthongs (FACE, SQUARE) were categorised as a single SA phonological vowel category /eː/. Therefore, they were analysed as the PAM SC assimilation type. PAM predicts that L2 learners will have great difficulty in discriminating the members of a SC contrast which is phonologically and phonetically assimilated to a single L1 category (Best and Tyler 2007: 29). However, the English (FACE, SQUARE) contrast was phonologically assimilated to the SA /eː/ vowel, but phonetically both members are equally deviant from that L1 category. Also, the participants were predicted to perceive the phonetic difference between the two English diphthongs, which is mostly in the off-glide part. English FACE is a rising diphthong whereas SQUARE is a centring one. As a result, it was predicted that the participants would have no difficulty in discriminating the members of this contrast since they perceive enough phonetic difference between these vowels. The results of the present study strongly supported this prediction. Apparently, the participants were able to perceive this difference between FACE and SQUARE vowels which had excellent identification and discrimination comparable to that of NE; and this contradicts PAM prediction for the SC assimilation type.

In the previous chapter, the mapping of the English (FACE, SQUARE) diphthongs into the SA front mid long vowel was analysed as an extension of the dialectal Arabic historical phonological process of vowel-glide assimilation into mid long vowels. In general, Arabs are aware of this phonological assimilation process of Standard Arabic vowel-glide sequences into mid long vowels since both Standard and dialectal Arabic coexists. Therefore, it can be argued that applying this process to the English diphthongal contrast does not affect the participants' ability to perceive the differences in the diphthongal projection between these vowels.

As suggested earlier, the PAT triggered participants to use their phonological vowel space, in which they were directed to use their dialect and not Standard Arabic. Therefore, the dialectal phonological form /eː/ was used in their categorisations rather than its standard equivalent vowel-glide sequence /aj/ (which was also included in the PAT). On the other hand, the identification task did not limit the participants' access to their phonological vowel space only. The participants were able to use their full phonological and phonetic vowel space, which arguably includes Standard Arabic categories. Therefore, the participants were able to access their Standard Arabic /aj/ form to identify and discriminate its phonetically equivalent English vowel FACE, as well as, their dialectal phonetic [e] and [ə] forms which are phonetically equivalent to the components of the English SQUARE vowel.

6.5.3 English Category Goodness (CG) contrasts

The results of the previous chapter revealed that two English vowel contrasts (KIT, DRESS) and (GOAT, THOUGHT) match the PAM description of the CG assimilation type. These contrasts were categorised as a single SA phonological category /i/ and /oː/, respectively. Within these contrasts, English KIT and THOUGHT were phonologically and phonetically equivalent to the L1 categories. On the other hand, the English DRESS and GOAT vowels were perceived as phonologically close to the L1 categories, but phonetically deviant. Generally, PAM predicts that the members of these contrasts will have from moderate to good discrimination, which is not as good as the discrimination of PAM TC contrasts (Best and Tyler 2007: 29). Additionally, PAM predicts a better identification of English KIT and THOUGHT vowels which were phonologically and phonetically closer to the SA vowels.

The results of the identification task supported the PAM predictions for the discrimination of both English CG contrasts, where the English (KIT, DRESS) contrast (26% of KIT was misidentified as DRESS, and 6% of DRESS was misidentified as KIT) and the (GOAT, THOUGHT) contrast (14% of GOAT was misidentified as THOUGHT, and 23% of THOUGHT was misidentified as GOAT) had moderate discrimination. As for the identification of the members of these contrasts, English THOUGHT had a slightly better correct identification than GOAT. However, English KIT vowel had significantly poorer identification than its counterpart DRESS, which contradicts the PAM prediction for better identification of the vowel closer to L1.

The very good identification of the English GOAT diphthong and the moderate discrimination of English (GOAT, THOUGHT) contrast can perhaps be attributed to the effect of the participants L1 phonetic schwa category. Since the identification task was able to trigger the participants to use their full phonological and phonetic vowel space, they were able to use their SA schwa vowel to identify the on-glide of GOAT diphthong, which is a schwa. Therefore, it was easier for the participants to identify the GOAT vowel and thus discriminate it from its counterpart THOUGHT. Thus, it can be argued that during the PAT the participants focused on their phonological categories whereas in the identification they were able to use their phonological and phonetic ones.

Again, the unexpectedly poor identification of the English KIT vowel can be attributed to the effect of the participants' L1/SA phonetic categories. In the PAT, the SA phonological /i/ vowel and its phonetic variants [e] and [ə] were included in the experiment, which confused the participants who were not used to or not aware of their L1 phonetic categories. Since these are phonologically related SA vowel categories, the participants distributed their categorisations of English (KIT, DRESS) vowels over the three SA categories. Unlike the SA (/aː/, [aːˤ]) contrast, the participants did not pick one of the three SA (/i/, [e], [ə]) categories as the canonical phonological form to bias their categorisations. A possible explanation for this difference might be that the acoustic phonetic difference between SA /aː/ and its pharyngealised allophone [aːˤ] is more prominent than the differences between SA /i/ and its phonetic variants [e] and [ə], see Chapter 4. Therefore, it can be argued that, unlike SA (/aː/, [aːˤ]), the participants were not able to detect the phonological form which represents the latter SA contrasts.

In order to overcome this confusion, the categorisations into the three SA categories were grouped together as one phonological category, i.e. SA /i/ vowel. The outcome of this grouping showed the expected pattern of significantly more categorisations of English KIT vowel as SA /i/ vowel, therefore, it was considered to be phonologically and phonetically close to that L1 category. However, these findings do not explain the poor identification of English KIT and its misidentification as English DRESS and not the other way round. A plausible explanation might be in the acoustic characteristics of these vowels. Perhaps the SA /i/ vowel is acoustically closer to English DRESS than KIT (this will be examined thoroughly in the following chapter). If this is the case, this suggestion can account for the better identification of English DRESS, as well as, the misidentification of English KIT as DRESS. To support this suggestion, closer examination of the discrimination results extracted as part of the identification results in section 6.4.4 reveals that KIT was only confused with DRESS; when it was presented with other English categories, it was correctly identified.

Additionally, word familiarity might have had an effect on the identification of English (KIT, DRESS) contrast. Familiar words have been reported to be easier to perceive than less familiar ones (Flege et al. 1998). As mentioned in the stimulus material (section 6.2), English 'bit' and 'fizz', and 'bet' and 'fez' words were used to represent this contrast. Unfortunately, the design of the present study did not include a measure to extract the participants' subjective ratings of word familiarity. Nevertheless, as an English language teacher in Syria, I can estimate that there is no great difference in the familiarity of the KIT and DRESS words, however, English 'fizz' and 'fez' might be less familiar to the participants than English 'bit' and 'bet'. Even if this was the case, the participants systematically identified KIT words as DRESS in both of these word pairs. Thus, it is unlikely that the familiarity of these words have affected the identification of this contrast.

Overall, Best et al. (2007: 29) argue that the participants are likely to develop new phonetic and phonological categories for the deviant L2 vowels DRESS[9] and GOAT, in this case, but they are not likely to develop new categories for the phonologically and phonetically equivalent to the L1 vowels, i.e. KIT and THOUGHT. Thus, the participants are expected to keep perceiving the latter English categories as exemplars of their L1 categories.

6.5.4 English Uncategorised- Categorised (UC) contrast

As for the English (FLEECE, NEAR) contrast, it was categorised as a single SA phonological vowel category /iː/. In the PAT, the FLEECE vowel was frequently categorised as the best exemplar of SA /iː/ with a high goodness of fit rating, thus, it was considered phonologically and phonetically equivalent to the L1 category. Also, English NEAR was categorised as the same L1 vowel /iː/, but with a lower goodness of fit rating and a lower percentage of categorisation. Thus, NEAR was analysed as phonologically close to the L1 category but phonetically deviant. The results of the PAT for this contrast matched the PAM description for the UC assimilation type, in which English FLEECE was the categorised vowel whereas NEAR was the uncategorised one. PAM predicts that the members of the UC contrast will have very good discrimination since the participants perceive the NEAR vowel as phonetically deviant from their L1 vowel,

[9] English DRESS is still considered the deviant from the L1 /i/ category till the production of English (KIT, DRESS) vowels are compared to that of SA /i/ to confirm which English vowel is acoustically closer to the L1 vowel.

and thus, they will perceive it as phonetically deviant from its counterpart the FLEECE vowel. The model also predicts very good identification of the deviant L2 vowel but a better identification for the phonologically and phonetically equivalent to the L1 category (Best and Tyler 2007: 28-29).

The results of the identification task supported the PAM predictions for the UC contrast. The members of this contrast were well discriminated and the English FLEECE vowel was significantly identified better than the deviant vowel NEAR. Similarly to the English SQUARE vowel, English NEAR is a diphthong that contains a schwa as its off-glide which makes it phonetically different from the L1 monophthongal vowel /iː/. Therefore, it can be argued that the participants were not able to perceive the schwa in NEAR as it constitutes a phonetic category in SA. Instead, the participants mapped English NEAR vowel to a long L1 category, which was close to the quality of its on-glide, i.e. SA /iː/ category. On the other hand, in the identification task, the participants were able to use their L1 phonetic categories to correctly identify and discriminate phonetically similar L2 categories. Thus, the phonetic L1 schwa category had a great effect on the identification of English NEAR and the discrimination of English (FLEECE, NEAR) contrast.

In sum, including SA allophonic phonetic categories in the PAT in the previous experiment had a great effect on understanding the categorisation patterns of the English vowels, as well as on the analysis of the results of the identification task. Best et al (2007: 19) argued that L2 perception is affected systematically not only by the L1 phonological system but also by its phonotactic, allophonic, and coarticulatory forms. Therefore, the effect of L1 phonological and phonetic categories (vocalic or consonantal) should be taken into consideration in any study examining L2 perception and/or production. Additionally, it is important to decide which level of analysis to aim for in an L2 study, i.e. phonological or both phonological and phonetic levels. This decision will have a great effect on the kind of perceptual test to be used in order to trigger the intended level of analysis.

6.6 Summary

To summarise, the results of the identification task showed that the members of the TC English (TRAP, BATH) and (LOT, STRUT) contrasts had excellent identification and discrimination by SA participants. Also, the members of the SC English (FACE, SQUARE) contrast which were rated as equally phonetically deviant from the L1 category had excellent identification and discrimination. As for the CG (KIT, DRESS) and (GOAT, THOUGHT), and the UC (FLEECE, NEAR) English contrasts, they had from moderate to very good discrimination and a better identification of the L2 category which is phonetically deviant from the L1 categories, except English KIT, which was confused with its counterpart DRESS vowel.

The identification task aimed to answer the following research questions as follows:

RQ1: How do SA participants perceive the English vowels, in terms of the L2 phonemic vowel inventory, compared to NE participants?

RQ2: Which of these English vowel contrasts are difficult to discriminate by the SA participants?

Compared to the NE participants, who had excellent identification and discrimination of the six English contrasts under investigation, the SA participants were best in identifying and discriminating the TC contrasts, as predicted by PAM. On the other hand, the CG contrasts were the most difficult to discriminate and the members which were deviant from the L1 categories were the most difficult to identify. Similarly, the uncategorised member in the UC contrast was difficult to identify and discriminate, but not as hard as the deviant members of the CG contrasts. Therefore, the English contrasts can be ordered as follows in terms of their perceived difficulty by the SA participants: TC, SC, UC, and CG.

RQ3: Is this difficulty related to quality or quantity overlap?

In general, English tense vs. lax vowels were mapped into SA long vs. short vowels, which is what would be expected because of the phonological role of length in Arabic (Munro 1993). As for English diphthongs, the participants focused on one of the diphthongal components and mapped that into its closest SA phonological long vowel. In chapter 4, the SA vowel system was analysed as a five-quality vowel system with a phonetic schwa category, whereas English has a larger vowel system, in which quality has the main role in the phonological distinctions of all of its vowels (Deterding 1997, Wells 1986). Therefore, it can be argued that the qualitative overlap had a greater effect than the quantitative one on the identification and discrimination of English vowels by SA participants. This suggests that SA participants had greater difficulty to perceive the qualitative differences between the English vowels than the quantitative ones. For example, the SA participants might have been able to perceive a durational difference between the members of the English CG and UC contrasts rather than (or greater than) a qualitative one. In particular, the members of these English contrasts occupy a single SA phonological category area in the vowel space, i.e. /i/, /oː/, and /iː/. Thus, any two English categories within a particular SA space are predicted to be perceived with overlapping qualities.

RQ4: To what extent can the predictions of PAM assimilation categories account for the results of the present study?

On the whole, the PAM predictions can account for the results of the identification and discrimination patterns of the present study if amended particularly for the group of learners under examination. The analysis of these results revealed that the PAM assimilation categories and their predictions cannot be directly applied to a particular group of foreign or second language learners. Full knowledge of the participants' L1 background including its phonological, phonetic, phonotactic, and coarticulation patterns, might have an effect on the interpretation of PAM categories and their resulted predictions. Bearing in mind these language specific characteristics, PAM can be a useful model to apply its testable categories as a starting point to identify and to account for the perceptual difficulties of non-native language learners.

In order to complement the perceptual analysis of the six English vowel contrasts under investigation, the following chapter presents a comparative acoustic analysis which was conducted to examine the production of these English vowels by the same SA participants, compared to NE productions, and to the productions of their L1 equivalent vowels. The results of the PAT and the present study showed that some of the English vowels were analysed as phonologically and phonetically equivalent to SA vowel categories, while others were analysed as phonologically similar but phonetically deviant. This perceptual difference between the two groups of vowels entails parallel production patterns, which means that a vowel which is phonologically and phonetically equivalent to an L1 vowel is expected to be produced acoustically in a more native-like manner than a phonetically deviant vowel.

Importantly, excellent perception of an English vowel need not imply that it is acoustically produced as identical to NE, in particular, because none of the English vowels was given an identical goodness of fit rating in the PAT. In the following chapter, the production patterns of the six English vowel contrasts under investigation will be examined to determine whether they are produced with Distinct acoustic Parameters (DP) or Overlapping Parameters (OP); and whether they are Native-like (NL) or Non-Native-Like (NNL).

7 Production Experiment

In this chapter, the participants' L2 productions of the six English vowel contrasts under investigation will be examined in terms of quality (F1/F2) and quantity (duration). Similarly to the production experiment of the pilot study, L2 productions of these vowels will be compared to their L1 equivalent productions and to NE productions of the same vowels. The present study extends the methods used by Yang (1996) for acoustic comparison between the vowels of two different languages to examine the same research questions of the pilot, following similar methods and procedures. However, based on the results of the previous PAT and the identification task, specific hypotheses and predictions about the L2 productions of the English vowel contrasts were examined. To review, the production experiment explores the following research questions:

RQ1: What are the qualitative (F1/F2) and quantitative (duration) acoustic characteristics of the participants' L2 English productions compared to NE and SA?

RQ2: What is the effect of the participants' L1 phonological vs. phonetic vowel space on their L2 productions?

RQ3: After reducing the non-linguistic effects of gender and language using LOBANOV statistical normalisation procedure, which acoustic measurements do reflect phonetic differences between English, SA, and/or L2 vowels, qualitative and quantitative?

RQ4: How do PAM perceptual predictions of the different assimilation types reflect on production?

This chapter is structured as follows: first, the hypotheses which were based on the results of the previous chapters are presented in section 7.1. Second, the methods used for this experiment are presented in section 7.2. The data analysis procedures are presented in section 7.3. After that, the results are presented for each English vowel contrast in section 7.4. Then, the results are discussed within PAM perceptual assimilation categories in section 7.5. Finally, the results of the present study are summarised in section 7.6.

7.1 Hypotheses

The production experiment provides a description of the acoustic characteristics of the SA participants' L2, which is linked to their perceptual patterns. The analysis of the results includes comparing the participants' L2 productions to their L1 equivalent productions and to NE productions. Based on the results of the PAT and the identification task, the following hypotheses were formulated to explore the productions of the six English vowel contrasts under investigation:

- In the PAT, the members of the English (TRAP, BATH) contrast were categorised with high goodness of fit ratings, mostly as the SA (/aː/, [aːˤ]) contrast. Also, the members of this English contrast received high correct identification and discrimination. Therefore, this contrast was analysed as the PAM TC assimilation type. In the present study, the L2 productions of these English vowels are predicted to be produced with distinct acoustic parameters (DP) since the participants had no difficulty discriminating them. However, those parameters might not exactly match NE productions (NNL Non Native-Like) given that SA participants did not give an identical goodness of fit rating to any of these vowels, i.e. (7/7). As a result, L2 productions are expected to be closer to their L1 equivalents or half way between NE and L1 productions.
- The English (LOT, STRUT) vowels were categorised mostly as SA ([o], /a/) vowel categories, with relatively low goodness of fit ratings. This contrast yielded excellent discrimination and high correct identification of English STRUT vowel. However, the English LOT vowel received moderate identification as it was misidentified as other English long vowels GOAT and THOUGHT. This contrast was also analysed as PAM TC assimilation type, and it is predicted to be produced with distinct acoustic parameters (DP). Similar to the previous contrast, these vowels are expected to be NNL and to be produced with acoustic parameters that are half way between NE and L1 vowels since the goodness of fit ratings were low.
- As for the English diphthongal contrast (FACE, SQUARE), it was analysed as PAM SC assimilation type as both vowels were categorised into a single SA category /eː/, with relatively high but not identical goodness of fit ratings (5/7).

Although this was a SC contrast, the participants were able to highly and correctly discriminate and identify the members of this contrast. Thus, the two vowels are predicted to be produced, similarly to the previous contrasts, with DP but NNL productions. Since, the participants were able to detect the phonetic differences between these vowels; they are predicted to attempt to produce these vowels with diphthongal trajectories. However, those trajectories might not match NE productions, particularly for the SQUARE vowel which ends with a schwa quality.

- The English (KIT, DRESS) vowels were analysed as PAM CG assimilation type since they were categorised as a single SA phonological category /i/ (including its phonetic variants [e] and [ə]), with KIT being closer to that category in the number of responses and the goodness of fit ratings. However, none of these vowels were given a high rate of similarity to the L1 category. Unexpectedly, the English DRESS vowel was highly correctly identified whereas KIT was moderately correctly identified, and both received good discrimination. The poor identification of English KIT and misidentification of the DRESS vowel was argued to be a result of acoustic similarity between English DRESS and the SA /i/ vowel. Therefore, both English vowels are predicted to be produced with overlapping acoustic parameters (OP) and NNL productions, which are expected to overlap with their L1 /i/ acoustic parameters.
- Similar to the previous contrast, English (GOAT, THOUGHT) vowels were analysed as a CG contrast, in which both were categorised as a single SA category /oː/, with THOUGHT being closer to that category. In the identification experiment, both English vowels received very good identification and discrimination. In production, these vowels are predicted to be produced with OP and NNL productions. L2 English THOUGHT productions are expected to be closer to L1 productions since it received a high similarity rate (6/7), i.e. NL productions, whereas the GOAT vowel is expected to be produced with a back quality that is half way between NE and L1 productions as it received less rating (4/7) and it was misidentified as other English back vowels, i.e. NNL productions.

- Finally, the English (FLEECE, NEAR) contrast was analysed as PAM UC assimilation type, in which FLEECE was the categorised vowel as it had high similarity with its L1 equivalent vowel /iː/ (6/7) and it was highly accurately identified. Therefore, it is predicted to be produced NL if NE FLEECE and L1 /iː/ vowel were acoustically similar. English NEAR vowel received less rating of similarity (5/7), and less correct identification. Additionally, NEAR was misidentified as the English FLEECE vowel. Thus, it is predicted to have NNL L2 productions, with little or no diphthongal trajectory. These vowels are predicted to be produced with OP which is close to the English FLEECE vowel.

Table 7-1:The predicted production patterns of the six English vowel contrasts compared to NE productions (Native-like (NL) or Non-native like (NNL)), and the production patterns of the members of each contrast (Distinct acoustic parameters (DP), i.e. F1/F2 and duration, or Overlapping acoustic parameters (OP))

English contrast	PAM type	Compared to NE	Production
(TRAP, BATH)	TC	(NNL, NNL)	DP
(LOT, STRUT)	TC	(NNL, NNL)	DP
(FACE, SQUARE)	SC	(NNL, NNL)	DP
(KIT, DRESS)	CG	(NNL, NNL)	OP
(GOAT, THOUGHT)	CG	(NNL, **NL**)	OP
(FLEECE, NEAR)	UC	(**NL**, NNL)	OP

Table 7-1 above summarises the hypotheses for the production experiment, showing the predicted production patterns of the six vowel contrasts under investigation compared to NE productions. Generally, L2 productions (DP/OP) of these vowels were predicted based on the PAT and identification results in the previous chapters 5 & 6, whereas the comparison to NE productions (NL/NNL) was predicted based on the goodness of fit ratings in the PAT in chapter 5. If a particular contrast is predicted to be produced with DP or OP, both members of this contrast can be NL or NNL, or one NL and the other NNL. In sum, the TC and the SC contrasts are predicted to be produced with DP, which means that the members of these contrasts should be produced with different F1/F2 and duration measurements, i.e. significant differences are predicted between the members of the TC and SC contrasts in terms of F1, F2, or duration. On the other hand, the CG and the UC contrasts are predicted to be produced with OP, which means that no significant differences are expected between the members of these contrasts in terms of F1, F2, or duration. Additionally, none of these vowels is expected to exactly match NE productions since they were not given identical similarity ratings to their L1 equivalent categories. However, a high goodness of fit rating of a particular English vowel (e.g. THOUGHT and FLEECE vowels) with a SA category suggests that NE, L2, and L1 productions will be produced with relatively similar acoustic characteristics.

7.2 Methodology

The recording procedures of the SA participants were the same as those followed in Chapter 4; for details see section 4.1. Also, the recording procedures for the NE participants were similar to those followed in the recordings of the NE speakers in the previous experiments; for details see section 3.1.2.

For technical reasons, five of the ten SA females and 10 SA males from the previous experiments were recorded for the production experiment; for more details see section 5.2. As mentioned earlier, the participants were asked to rate their language skills on a seven point scale. On average, the rate for their speaking skill was the least (4.8/7) and this suggested that the SA participants would have difficulty producing what matches NE productions. As a control group, the same five female and four male NE participants from the identification experiment were also recorded, for more information about the NE participants see section 6.2.

The English stimuli and the instructions were the same for both the SA and the NE participants. The SSBE stimuli were the same as the ones used in the pilot production experiment; for full details see section 3.3.4.1. The main focus of the present study was to examine the productions of the six English vowel contrasts under investigation, and the rest of the English vowels were included as distracters and for further research in the future. Similarly to the pilot study, the best three productions of each vowel were extracted and analysed for each participant. The analysis covered the productions of the English vocalic contrasts (TRAP, BATH), (FACE, SQUARE), (KIT, DRESS), (LOT, STRUT), (GOAT, THOUGHT) and (NEAR, FLEECE). Overall, 3 repetitions × 12 vowels × 24 participants = 864 items were analysed for this experiment.

The SA stimuli used for this experiment were the same as the ones used in the analysis of the SA vowel system in Chapter 4; for more details see section 4.1. Similarly to the SSBE stimuli, the best three productions of each vowel for each participant were extracted and analysed. Unlike Chapter 4, here the effect of L1 was examined on the SA participants' L2 productions of only the six English vowel contrasts under investigation. Therefore, only the analysis of SA (/aː/, [aːˤ], /eː/, /ɪ/, [e], [ə], [o], /a/, /oː/, and /iː/) vowels, which were categorised into those English vowels in the PAT, was included in this study. Together, SA stimuli included 3 repetitions × 10 vowels × 15 participants = 450 items.

The procedures of the production experiment of the present study were similar to those followed in the pilot study (section 3.3.5), except for the presentation of the material to the participants. Unlike the production experiment of the pilot study, the English and SA stimulus material were presented to the participants using flash cards rather than a PPT slide-show. The flash cards were used in order to get to get clear sound files without any noise from the computer. Each of these cards contained two stimulus phrases. The cards were randomised and then presented to the participants. The participants were directed to read the phrases on each card naturally and at a normal speed.

7.3 Data Analysis

The analysis procedures for this study follow the ones used in the pilot study (section 3.3.6), in which vowel duration, and F1 and F2 mid-point frequency measurements were extracted. Additionally, vowel duration and the frequency measurements were normalised following the same methods used in the analysis of the SA vowel system in chapter 4 (section 4.2).

Using SPSS, three linear mixed model tests were run to examine the differences between NE productions of a particular vowel compared to SA participants' L2 productions of the same English vowel, and both NE and L2 productions were compared to SA participants' L1 productions of the closest SA equivalent vowel. The normalised vowel duration, normalised F1, and normalised F2 measurements were used as dependent variables for each vowel. For each test, 'language' with three levels: NE, SA (L1), and SA (L2), and 'sex' with two levels (M/F) were used as fixed factors, whereas 'participant' (24 participants) was used as a random factor. Since the L1 and L2 results are from the same SA speakers, the SA participants' codes were used for their L1 and L2 productions. Accordingly, the linear mixed model took into consideration that L1 and L2 productions are for the same SA speakers, compared to NE productions. Similarly, the repeated measures of the three points of diphthongal measurements were given the same code for the vowel type which means that the linear mixed model took into account that they are repeated measures of the same vowel by the same speaker.

7.4 Results

To have a comparable presentation of the results of the present study to those of the PAT and the identification task, the results are presented for each of the six English vowel contrasts which were analysed and ordered in terms of PAM perceptual categories. As illustrated earlier, the productions of the members of these contrasts will be compared between NE and L2 participants' productions and their L1 equivalent productions.

7.4.1 The Production of the English (TRAP, BATH) contrast

7.4.1.1 TRAP

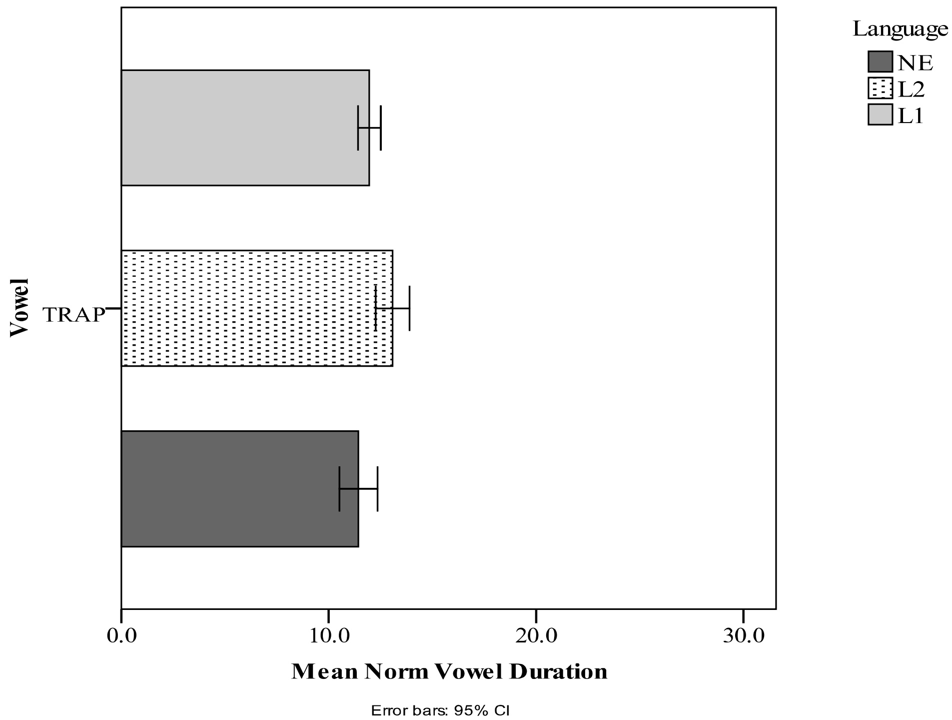

Figure 7-1: Mean normalised vowel duration of SA participants' L2 productions of English TRAP vowel compared to NE productions and their L1/SA equivalent vowel /aː/.

Figure 7-1 above presents mean normalised vowel duration of the English TRAP vowel as produced by NE and L2 participants compared to its SA equivalent vowel /aː/. As can be seen, L2 productions of English TRAP are slightly longer in duration than both NE and L1 productions. A liner mixed model effects was run to examine language and sex effects on normalised vowel duration differences. The results are summarised in Table 7-2 below, which shows that language had a significant effect on these differences, whereas sex and language*sex interaction had no significant effect. The results of the Bonferroni post hoc test showed that L2 duration was significantly longer than both NE ($p = .032$) and L1 ($p = .054$) productions, which were similar ($p = 1$).

Table 7-2: Summary of the results of the linear mixed model tests which were run to examine language, sex, and language*sex interaction effects on the differences in normalised vowel duration, normalised F1, and normalised F2.

Vowel	Fixed effects	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
TRAP	Language	$F(2,108) = 4.38$, **$p = .015$***	$F(2,108) = 1.28$, $p = .281$	$F(2,108) = 9.05$, **$p < .000$***
	Sex	$F(1,108) = 0.66$, $p = .418$	$F(1,108) = 4.71$, **$p = .032$***	$F(1,108) = 4.84$, **$p = .030$***
	Language*sex	$F(2,108) = 2.06$, $p = .132$	$F(2,108) = 2.09$, $p = .129$	$F(2,108) = 3.67$, **$p = .029$***

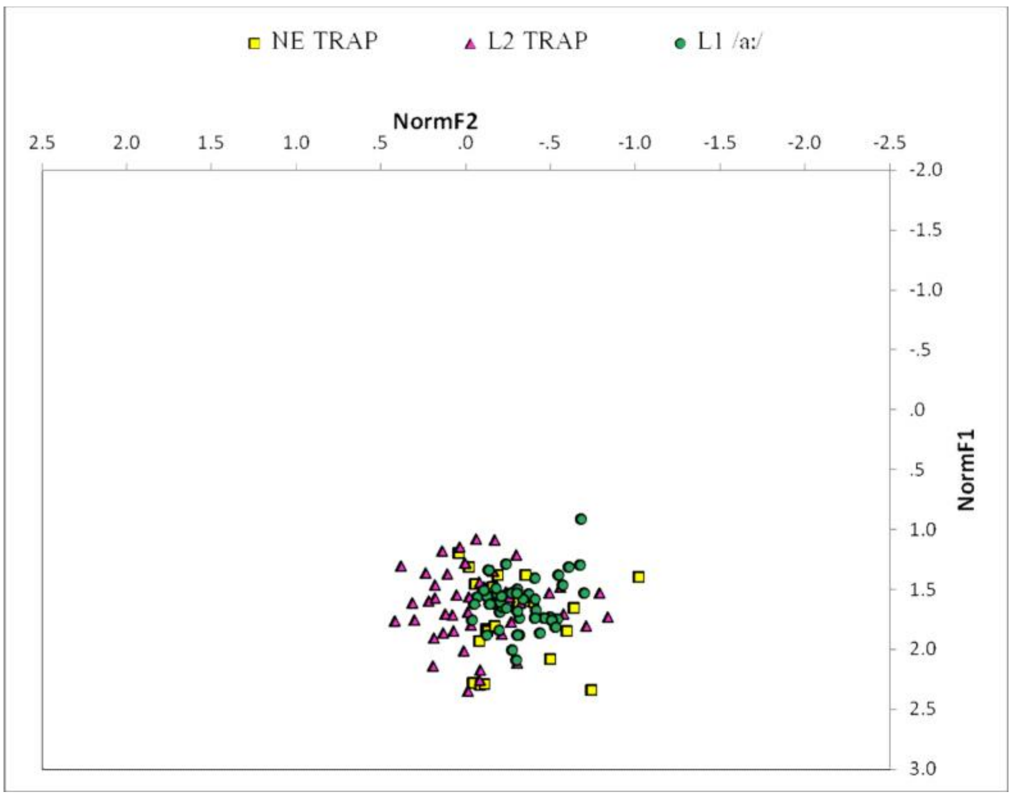

Figure 7-2: Mean normalised F1 and F2 values (following LOBANOV procedure) of SA participants' L2 productions of English TRAP vowel compared to NE and L1 productions.

In terms of vowel quality, Figure 7-2 above presents normalised F1 and F2 frequencies of SA participants' L2 productions of English TRAP compared to NE and L1 productions. In general, L2 productions are more dispersed than both NE and L1 productions. As can be seen, the productions of the three groups overlap; however, L2 productions appear fronter in the vowel space. Two linear mixed models were run to examine F1 and F2 differences between the three language groups. Table 7-2 above shows that language had no effect on F1 differences of English TRAP. However, sex effect was significant, while language*sex interaction effect was not. As for F2 differences, the effects of the three fixed factors were significant. Further examination of the Bonferroni post hoc test revealed that L2 productions were significantly fronter than both NE ($P = .009$) and L1 ($p < .000$), which were similar ($p = 1$). Additionally, NE and L2 male productions were fronter than females', whereas L1 female productions were fronter.

These results suggest that L2 productions of English TRAP vowel were not native-like (NNL) since they were significantly different in terms of quantity and quality, which supports the hypothesis for this vowel. Assignment of a high but not identical goodness of fit rating in the PAT is reflected in the production patterns of this vowel, which was produced in close proximity to both NE and L1 productions. Although NE productions of TRAP and L1 /aː/ vowels were very similar in this study, in the identification study, they were not perceived as parallel vowels, which resulted in a significantly different L2 productions which might have intended to match NE productions. Thus, it can be argued that having NNL TRAP productions cannot be attributed to the effect of L1 /aː/ vowel since it is very similar to NE TRAP production, rather it may be the result of the participants' perceptual patterns of the TRAP vowel.

7.4.1.2 BATH

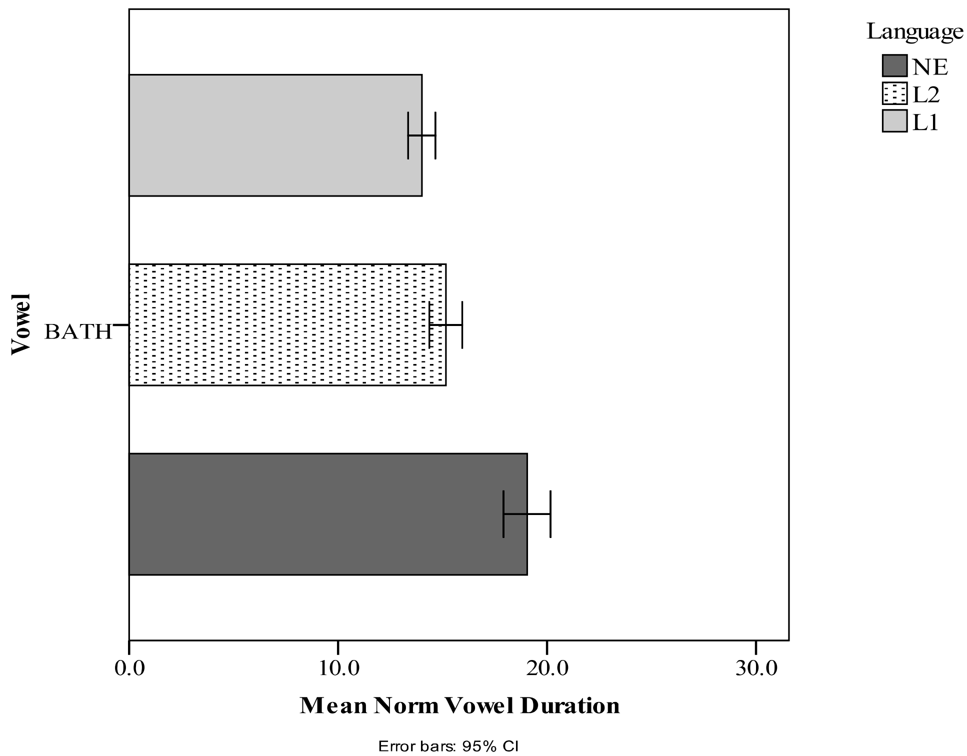

Figure 7-3: Mean normalised vowel duration of SA participants' L2 productions of English BATH vowel compared to NE productions and their L1/SA equivalent vowel [aːˤ].

As for L2 productions of the English BATH vowel, Figure 7-3 above shows that NE productions are longer than both L2 and L1 productions. Similar to English TRAP, a linear mixed model effect was run to examine language and sex effects on normalised vowel duration differences. Table 7-3 below shows similar effects to TRAP, in which only language had a significant effect. The results of the Bonferroni post hoc tests showed that NE productions were significantly longer than both L2 ($p < .000$) and L1 ($p < .000$) productions, which were similar ($p = .152$).

Table 7-3: Summary of the results of the linear mixed model tests for the BATH vowel

Vowel	Fixed effects	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
BATH	Language	F (2,107) = 31.29, **p < .000***	F (2,107) = 3.44, **p = .035***	F (2,107) = .45, p = .639
	Sex	F (1,107) = .151, p = .699	F (1,107) = .70, p = .405	F (1,107) = 4.52, **p = .036***
	Language*sex	F (2,107) = .503, p = .606	F (2,107) = 3, **p = .054***	F (1,107) = 1.55, p = .217

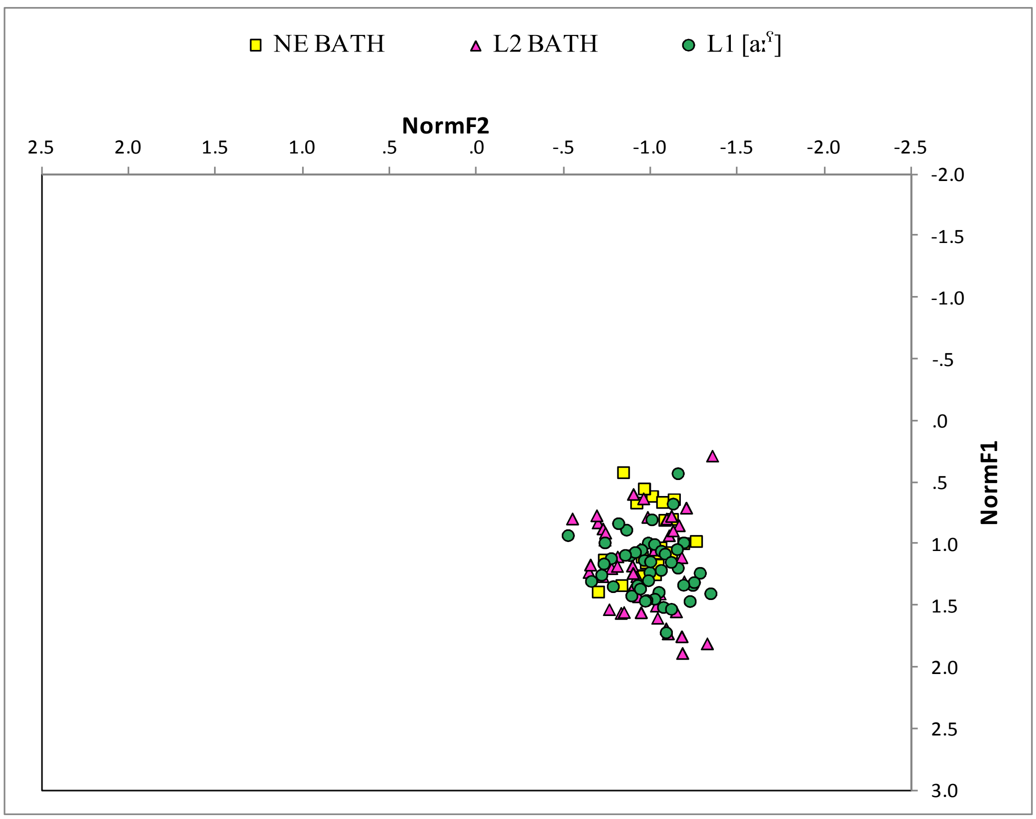

Figure 7-4: Mean normalised F1 and F2 values (following LOBANOV procedure) of SA participants' L2 productions of English BATH vowel compared to NE and L1 productions.

In terms of quality, Figure 7-4 above shows a greater overlap for the BATH vowel among the three groups and a greater clustering within each group than was seen for the TRAP vowel. However, NE productions appear higher than their L2 and L1 equivalents, which is supported by the results of the linear mixed model test. The results in Table 7-3 above show that Language had a significant effect on F1 differences, but not on F2 differences. Further examination of F1 values showed that L2 productions were not significantly different from NE productions ($p = .193$) nor from L1 productions ($p = 1$), which were significantly different from each other ($p = .031$). Additionally, the sex effect was not significant for F1, but language*sex interaction was marginally significant. Additionally, NE and L1 male productions were higher than females', whereas L2 female productions were higher than L2 males'. As for F2, only the sex effect was significant, while language*sex interaction was not.

The results for L2 productions of the English BATH vowel revealed that those productions matched NE BATH in terms of quality, but not in quantity, which supports the hypothesis for this vowel, i.e. NNL production. Producing L2 BATH with shorter durations can be attributed to the fact that Arabic is a rhotic language, and using the word 'hard' to extract this vowel has prompted, at least in some cases, a rhotic production with a shorter vowel duration. Another explanation might be that the target variety for the SA learners is American English rather than SSBE, which prompted different vocalic realisations from those produced by the NE participants. The L1 effect on L2 productions of BATH is more prominent than on TRAP. Particularly, L1 and L2 productions of BATH were not significantly different from each other in terms of quantity and quality. The latter finding suggests that learners were using a phonetic allophonic category [aːˤ], which they might not be aware of, to produce a phonetically similar L2 category.

7.4.1.3 L2 (TRAP, BATH)

The hypothesis for L2 productions of English (TRAP, BATH) contrast was that the members of this contrast were predicted to be produced with distinct acoustic parameters DP, i.e. a difference in at least one of the acoustic parameters: vowel duration, F1, or F2 would count two vowels as different. In order to examine this hypothesis, three linear mixed models were run to examine the differences between L2 productions of the two vowels. The results of these tests are presented in Table 7-4 below, which strongly support this hypothesis. L2 productions of English (TRAP, BATH) are highly significantly different from each other in terms of the three dependent variables, which can be seen in Figure 7-5 and Figure 7-6 below.

Table 7-4: Summary of the results of the linear mixed model tests for L2 (TRAP, BATH) productions

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Vowel	F (1,86) = 12.82, **p = .001***	F (1,86) = 56.18, **p < .000***	F (1,86) = 234, **p < .000***
Sex	F (1,86) = .915, p = .341	F (1,86) = .002, p = .961	F (1,86) = 1.64, p = .203
Vowel*sex	F (1,86) = .127, p = .723	F (1,86) = 11.83, **p = .001***	F (1,86) = 1.41, p = .238

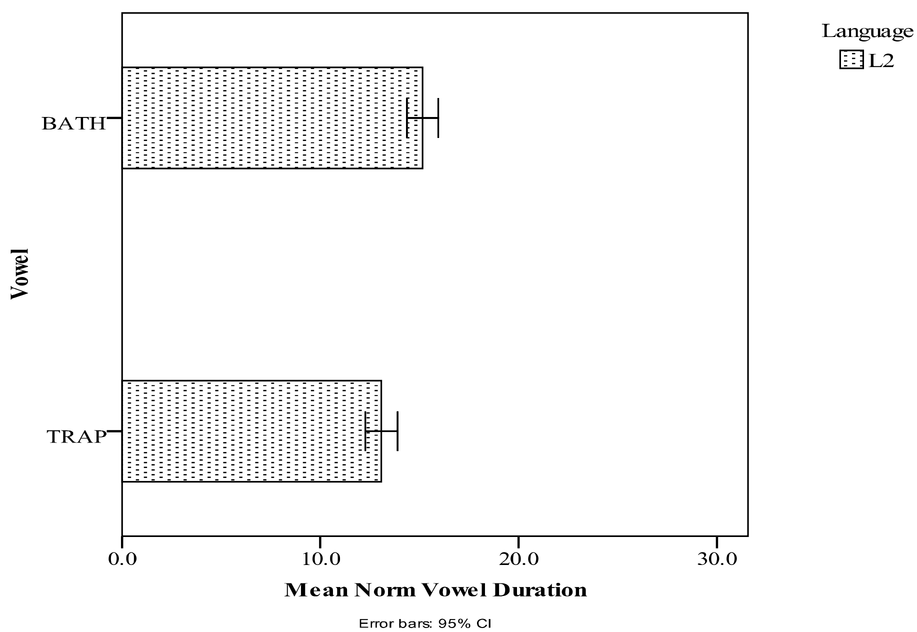

Figure 7-5: Mean normalised vowel duration of SA participants’ L2 productions of English (TRAP, BATH) contrast

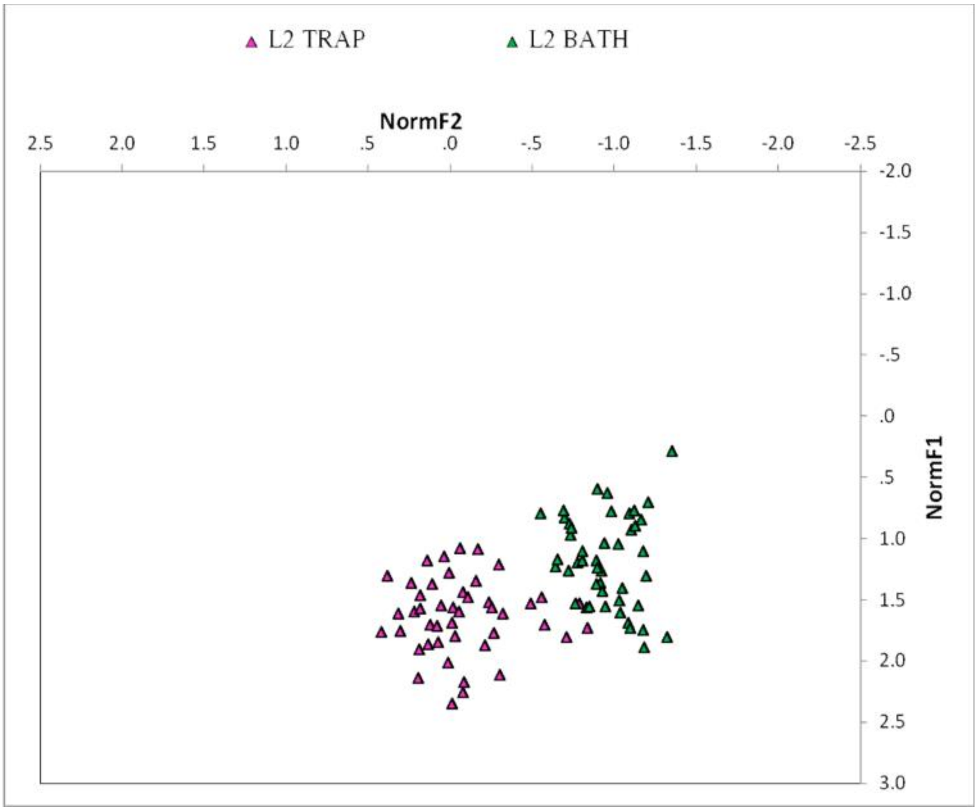

Figure 7-6: Mean normalised F1 and F2 values of SA participants' L2 productions of English (TRAP, BATH) contrast

In sum, SA speakers were able to produce the members of the English (TRAP, BATH) contrast with a great acoustic distance using an allophonic category /aːˤ/ from their L1. However, their L2 productions did not match NE productions though they were able to correctly identify and discriminate the two vowels in the previous experiment. This finding suggests that better identification of a particular L2 contrast leads to production of the same contrast with a greater acoustic distance, but not necessarily identical to NE productions.

7.4.2 The Production of the English (LOT, STRUT) contrast

7.4.2.1 LOT

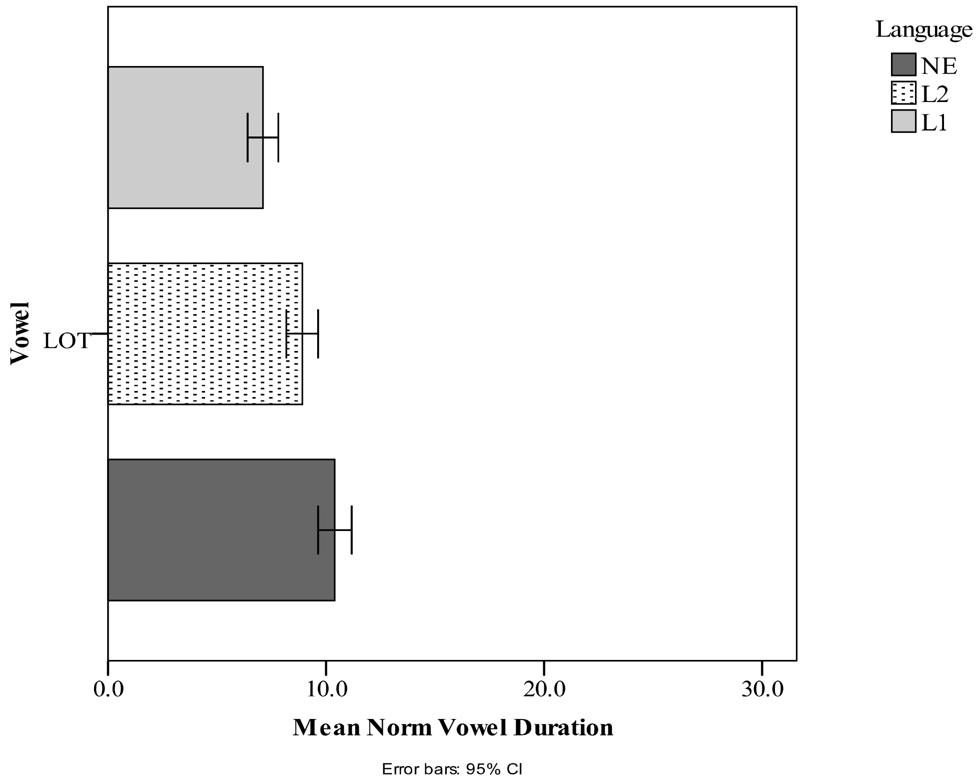

Figure 7-7: Mean normalised vowel duration of SA participants' L2 productions of English LOT vowel compared to NE productions and their L1/SA equivalent vowel [o].

Figure 7-7 above presents normalised vowel duration of L2 productions of the English LOT vowel compared to NE and L1 productions. As shown, NE productions are longer than L2 productions, which are also longer than their L1. The results of the linear mixed model given in Table 7-5 below confirm that language had a significant effect on durational differences. However, sex and language*sex interaction effects were not significant. The Bonferroni post hot tests showed that NE productions were significantly longer than L2 productions (p = .009), which were significantly longer than their L1 (p = .003), and the difference was greater between NE and L1 (p < .000). These finding suggest that the participants managed to produce English LOT vowel with a duration that is half way between their L1 equivalent vowel and the NE target.

Table 7-5: Summary of the results of the linear mixed model tests for LOT vowel

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Language	F (2,108) = 17.98, **p < .000***	F (2,108) = 53.54, **p < .000***	F (2,108) = 36.22, **p < .000***
Sex	F (1,108) = 1.38, p = .242	F (1,108) = 3.54, p = .062	F (1,108) = 1.13, p = .289
Language*sex	F (2,108) = 2.303, p = .105	F (2,108) = 4.47, **p = .014***	F (2,108) = 4.69, **p = .011***

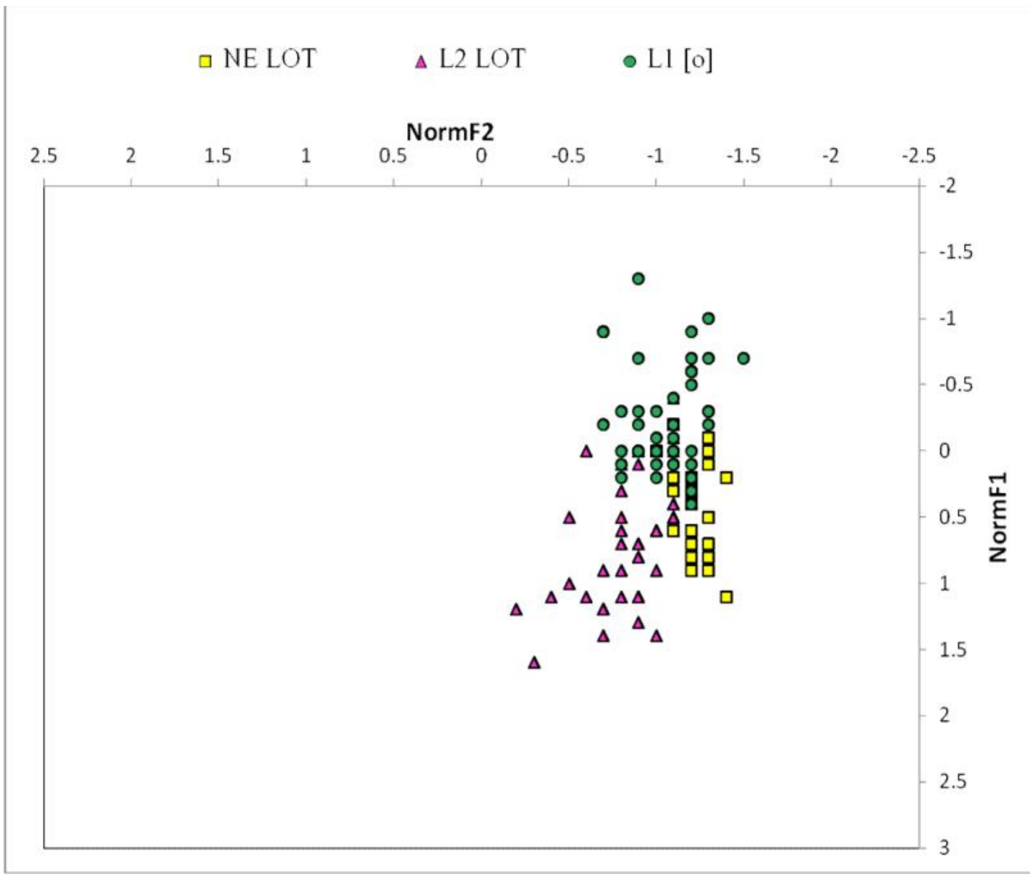

Figure 7-8: Mean normalised F1 and F2 values of SA participants' L2 productions of English LOT vowel compared to NE and L1 productions.

In terms of quality, Figure 7-8 shows that the productions of the three groups are close to each other, but there is no great overlap. The results in Table 7-5 show that language had a highly significant effect on F1 and F2 differences. Further examination of the Bonferroni post hoc tests showed that L2 productions do not differ from NE productions in F1 ($p = .109$), but they significantly differ in F2 ($p < .000$), in which L2 productions had a fronter quality. Additionally, L2 productions had significantly lower ($p < .000$) and fronter ($p < .000$) quality than their L1 productions. Table 7-5 also shows that the language*sex interaction effect is significant for F1 and F2, where L2 female participants had higher F1 and F2 values, which means that they were responsible for the lower and fronter L2 productions of LOT vowel.

The above results support the hypothesis for L2 productions of English LOT vowel, which was predicted to have NNL productions. The results of the present study showed that L2 productions were significantly shorter and fronter than NE productions. These results also reflect the moderate perception of this vowel in the previous chapters, in which it was given a low goodness of fit to the L1 category, and it received a 51% of correct identification. The effect of the L1 phonetic category [o] does not seem to be prevailing on L2 LOT productions. As mentioned in the previous chapter, SA participants associated English LOT with longer L1 back vowels in the PAT (32% as SA /oː/, see section 5.3). Thus, producing L2 LOT as significantly NNL was predicted.

7.4.2.2 STRUT

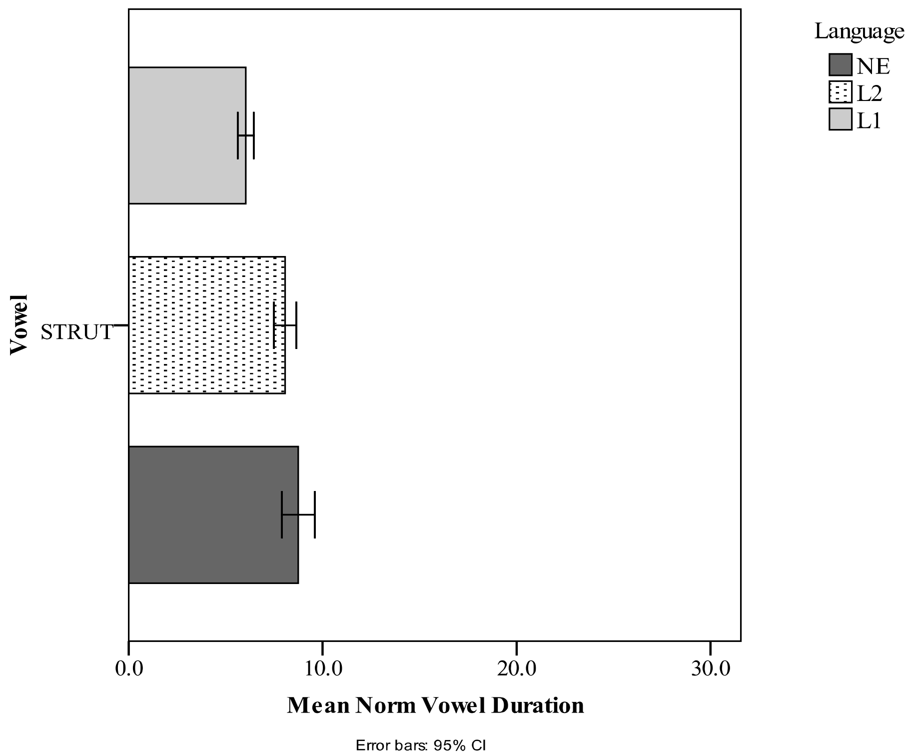

Figure 7-9: Mean normalised vowel duration of SA participants' L2 productions of English LOT vowel compared to NE productions and their L1/SA equivalent vowel /a/.

The results for the productions of the English STRUT vowel given in Table 7-6 below show that language had a significant effect on the durational differences. The Bonferroni post hoc tests revealed that L2 productions were not significantly different from NE in duration ($p = .140$), whereas L1 productions were significantly shorter than both NE ($p < .000$) and L2 productions ($p < .000$), which can be seen in Figure 7-9 above.

Table 7-6: Summary of the results of the linear mixed model tests for STRUT vowel

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Language	F (2,107) = 21.63, **p < .000***	F (2,108) = 12.53, **p < .000***	F (2,108) = 142, **p < .000***
Sex	F (1,108) = 2.98, p = .087	F (1,108) = .10.31, **p = .002***	F (1,108) = 19.88, **p < .000***
Language*sex	F (2,108) = 1.28, p = .281	F (2,108) = 10.11, **p < .000***	F (1,108) = 6.23, **p = .003***

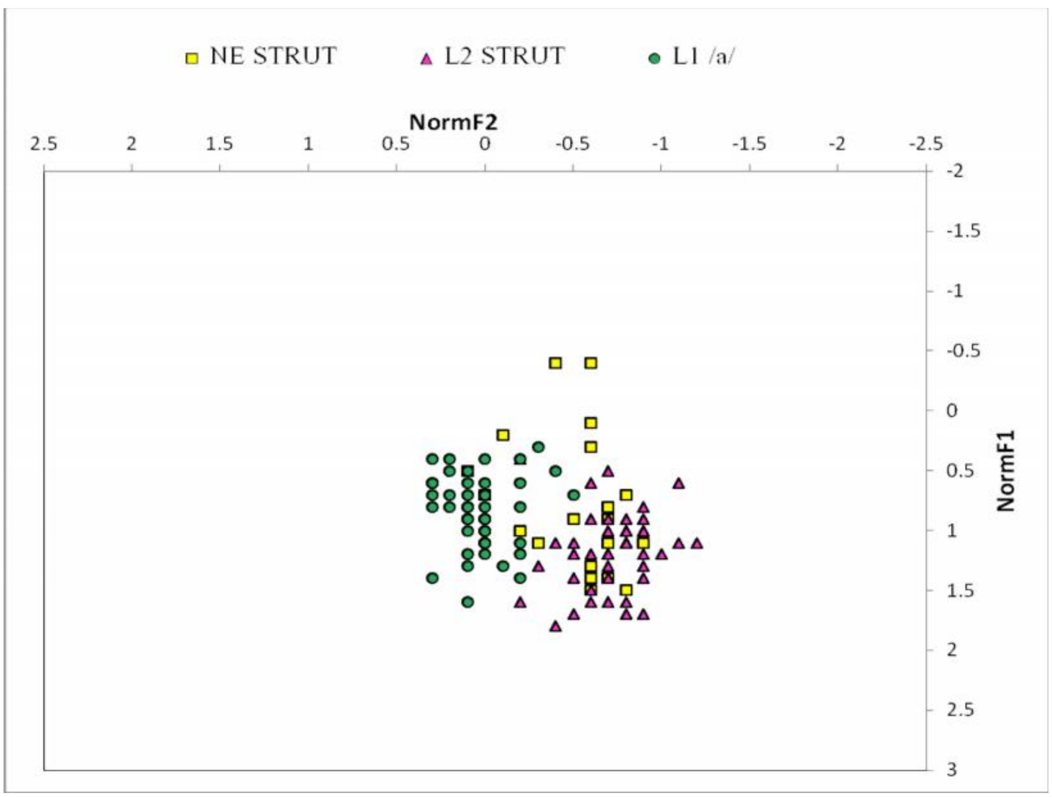

Figure 7-10: Mean normalised F1 and F2 values of SA participants' L2 productions of English STRUT vowel compared to NE and L1 productions.

Similarly to the previous vowel, Table 7-6 shows that language had a significant effect on the qualitative differences F1 and F2. The results of the Bonferroni post hoc tests showed that L2 productions were significantly lower than both NE (p < .000) and

L1 productions ($p < .000$). Also, L2 productions were backer than NE ($p = .003$) and L1 ($p < .000$) productions, which can be seen in Figure 7-10 above. Furthermore, NE and L1 productions were similar in height ($p = 1$), but L1 productions were significantly more centralised ($p< .000$). Like the LOT vowel, the language*sex interaction was significant for F1 and F2 differences, in which SA females' L2 productions were lower than their L1 and NE female productions; and both males' and females' L1 productions' were more centralised than their L2 and NE male and female productions.

Similarly to the previous vowel, the results support the hypothesis for L2 STRUT, which showed that it was NNL since it was significantly different from NE productions. Also, the effect of its closest L1 category /a/ was not prominent as L2 productions were significantly different from their L1. The perceptual patterns of this L2 vowel were reflected on its production. In the PAT, the STRUT vowel was categorised as the SA /a/ category, but it was also associated with its long vowel /aː/. Thus, like the LOT vowel, producing STRUT with a longer duration than its L1 equivalent was to be expected. One can argue that, though the participants' productions were close to those of NE in terms of quality, duration rather than quality might have had the greater role in the high correct identification (95%) of this vowel in the identification experiment.

7.4.2.3 L2 (LOT, STRUT)

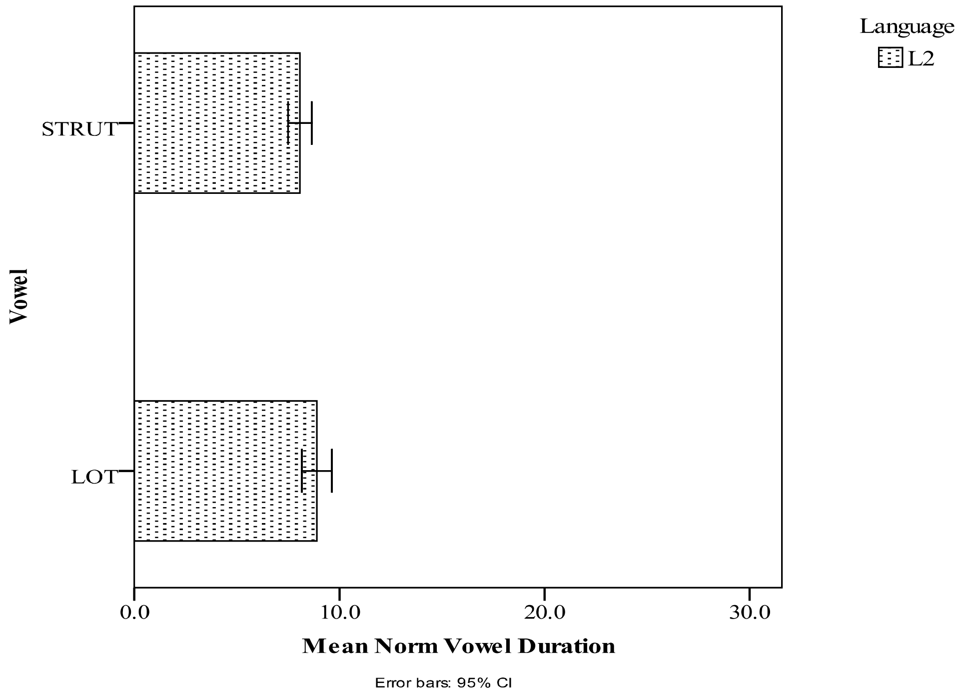

Figure 7-11: Mean normalised vowel duration of SA participants' L2 productions of English (LOT, STRUT) contrast

The hypothesis for L2 productions of the English (LOT, STRUT) contrast was that these vowels were predicted to be produced with distinct acoustic parameters DP. However, Figure 7-11 above shows that quantitatively both vowels were produced with similar durations, which is confirmed by the results in Table 7-7 below. On the other hand, the qualitative results in Table 7-7 below show that L2 productions of these vowels were significantly different. Similar to the previous TC contrast (TRAP, BATH), the members of this contrast were produced acoustically distinct in the vowel space, in which L2 STRUT was lower and more centralised than L2 LOT productions, which can be seen in Figure 7-12 below. Thus, the results of the present study confirm the hypothesis for the production of this contrast.

Table 7-7: Summary of the results of the linear mixed model tests for L2 (LOT, STRUT) productions

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Vowel	F (1,86) = 2.83, p = .96	F (1,86) = 30.75, **p < .000***	F (1,86) = 7.9, **p = .006***
Sex	F (1,86) = 9.7, **p = .002***	F (1,86) = .680, p = .412	F (1,86) = 10.7, **p = .002***
Vowel*sex	F (1,86) = .07, p = .780	F (1,86) = .593, p = .443	F (1,86) = .005, p = .942

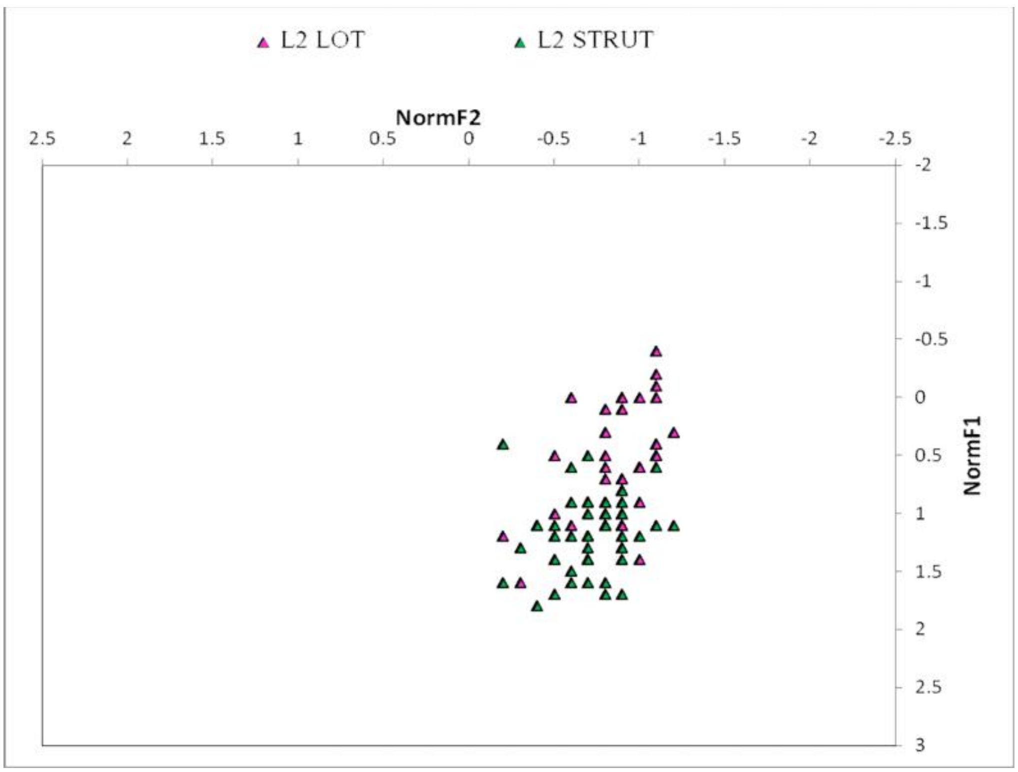

Figure 7-12: Mean normalised F1 and F2 values of SA participants' L2 productions of English (LOT, STRUT) contrast

Generally, SA participants managed to acoustically produce distinct English (LOT, STRUT) vowels, which would be expected from the excellent discrimination they received in the identification task. However, those vowels were not given a high rate of similarity to any SA vowel in the PAT, which yielded L2 productions that did not match NE nor L1 productions. The NNL production of the English LOT vowel can be attributed to its being associated, in the PAT, with a phonetic [o] rather than a phonological L1 category. Unlike the phonetic allophone [aːˤ], the acoustic properties of SA [o] are not prominent since its acoustic parameters are in the vicinity of another SA phonological category /u/. Thus, the SA [o] category might have been taken by the participants as its phonological variant /u/, which is acoustically different from the LOT vowel. Thus, the phonetic SA [o] category did not have a great effect on the L2 production of English LOT vowel. On the other hand, the NNL production of L2 STRUT vowel can be due to the lack of central vowels in SA.

7.4.3 The Production of the English (FACE, SQUARE) contrast

7.4.3.1 FACE

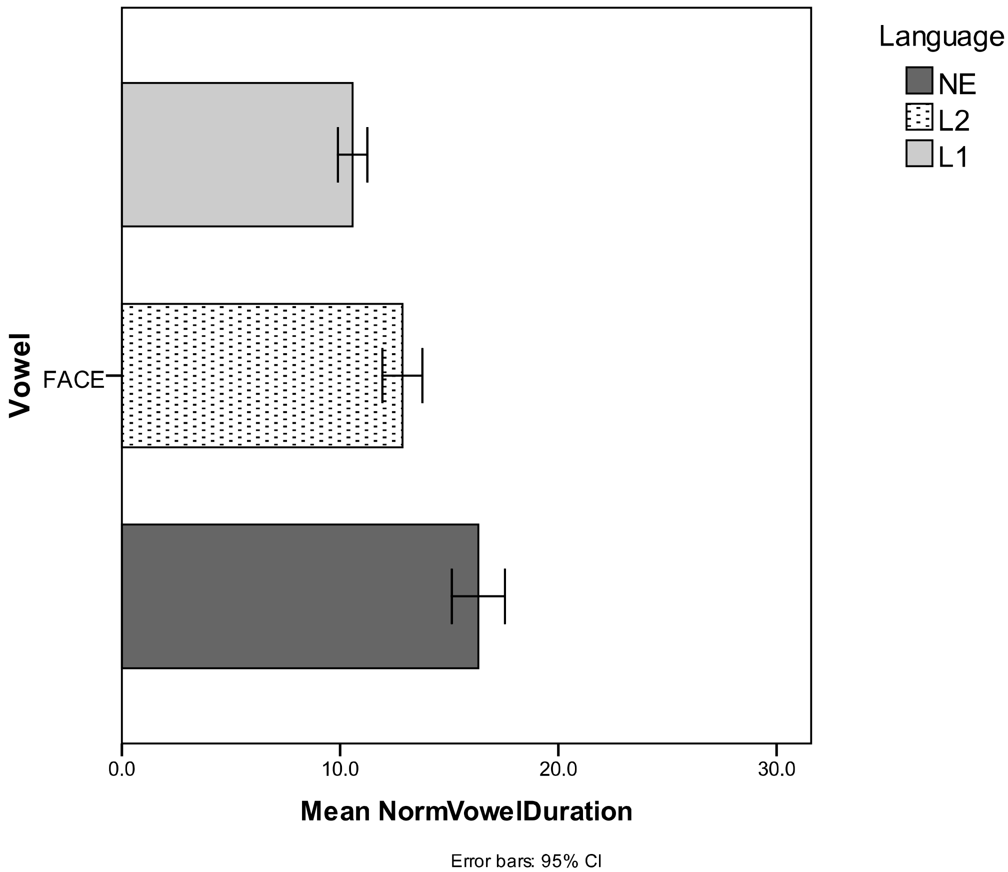

Figure 7-13: Mean normalised vowel duration of SA participants' L2 productions of English FACE vowel compared to NE productions and their L1/SA equivalent vowel /eː/.

The production of the L2 English FACE diphthong appears shorter than NE productions, but longer than their L1 equivalent productions of the SA /eː/ vowel, as shown in Figure 7-13. The results of the linear mixed models in Table 7-8 below show that language had a significant effect on the durational differences. The results of the Bonferroni post hoc test confirmed that the productions of the three groups significantly differ from each other, where L2 productions were significantly shorter than NE productions ($p < .000$), but longer than L1 productions ($p = .001$). Additionally, NE productions were significantly longer than L1 ($p < .000$).

Table 7-8: Summary of the results of the linear mixed model tests for FACE vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,108) = 36.14, **p < .000***	F (2,108) = 1.9, p = .141	F (2,108) = 5.08, **p = .008***
	50%		F (2,108) = 24.5, **p < .000***	F (2,108) = 17.1, **p < .000***
	75%		F (2,108) = 43, **p < .000***	F (2,108) = 53.5, **p < .000***
Sex	25%	F (1,108) = 1.24, p = .268	F (1,108) = .2.68, p = .104	F (1,108) = 4.9, **p = .028***
	50%		F (1,108) = .295, p = .588	F (1,108) = .426, p = .515
	75%		F (1,108) = .718, p = .399	F (1,108) = .22, p = .64
Language*sex	25%	F (2,108) = .642, p = .528	F (2,108) = 2.8, p = .063	F (2,108) = 1.77, p = .175
	50%		F (2,108) = 1.9, p = .153	F (2,108) = .035, p = .603
	75%		F (2,108) = 8.47, **p < .000***	F (2,108) = .503, p = .515

As for the qualitative differences, the normalised frequencies are presented in Figure 7-14 for NE, L2, and L1 productions, respectively. Unlike the previous vowels, the productions of this vowel are presented for each group separately in order to show the diphthongal trajectories from 25% through 50% to 75% of the vowel. Figure 7-14 shows that NE FACE productions exhibit a clear trajectory movement from a front mid vowel /e/ into a front high vowel /ɪ/, which constitute the on-glide and off-glide of the /eɪ/ diphthong. On the other hand, L2 productions do not show a clear trajectory movement, and the measurements across the three points of the vowel are more clustered than NE productions, which suggests a monophthongal L2 production. The L2 monophthongal FACE productions appear closer to their L1 productions, in which the measurements across the three points of the vowel overlap to form what would be expected in Arabic, i.e. a monophthongal vowel.

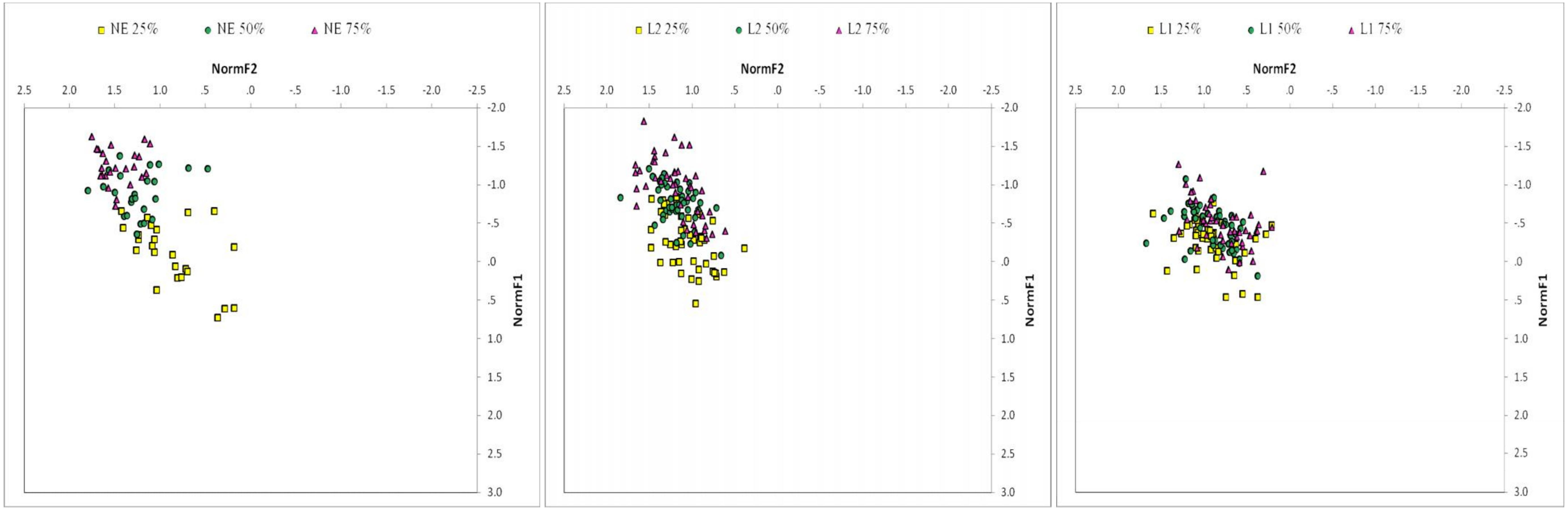

Figure 7-14: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English FACE vowel as produced by NE, L2, and L1 participants.

Table 7-8 above shows that language had a significant effect on F1 differences at the midpoint and towards the end of the vowel. The results of the post hoc tests revealed that L2 productions do not differ from NE midpoint F1 ($p = .07$). However, NE productions had a significantly higher off-glide than L2 productions ($p < .000$), which also suggests that the trajectory of the L2 productions was not clear and did not reach its off-glide target. As for L1 productions, they were significantly lower than NE ($p < .000$) and L2 productions ($p < .000$). In terms of F2/place, the language effect was significant throughout the vowel, in which L2 productions matched NE at midpoint ($p = 1$), but they were fronter at the beginning ($p = .018$) and backer at the ending of the vowel ($p < .000$). As would be predicted for the SA monophthong /eː/, it started from a place that was not significantly different from the NE beginning ($p = 1$), but slightly backer than L2 beginning ($p = .034$). Additionally, L1 productions were significantly backer than NE and L2 productions at the midpoint and towards the end of the vowel. This latter finding suggests that SA participants were trying to produce the FACE vowel with a trajectory, which means that the participants were trying to produce this phonetic detail which they were argued to be able to perceive (see section 6.1.). Thus, L2 FACE productions were produced with a small trajectory, but that did not match NE productions, which supports the hypothesis of the predicted NNL L2 productions of this vowel.

7.4.3.2 SQUARE

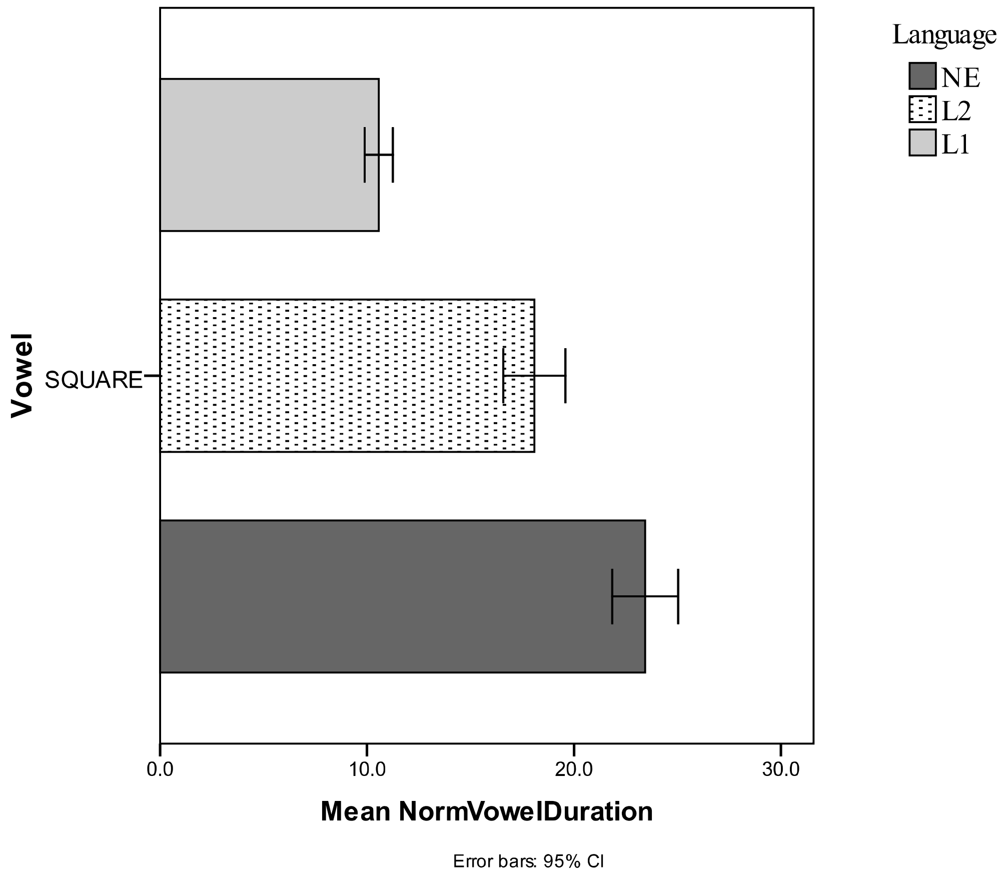

Figure 7-15: Mean normalised vowel duration of SA participants' L2 productions of English SQUARE vowel compared to NE productions and their L1/SA equivalent vowel /eː/.

Figure 7-15 above presents mean normalised vowel duration for the English SQUARE vowel comparing NE, L2, and L1 equivalent vowel /eː/ productions. The figure shows that NE productions appear longer in duration than L2 and L1 productions. Table 7-9 shows that language had a significant effect on durational differences, in which the three groups were highly significantly different from each other (p < .000). Thus, this vowel shows a similar pattern to English FACE, where L2 productions were produced with a duration that was longer than their L1 but shorter than NE productions.

Table 7-9: Summary of the results of the linear mixed model tests for SQUARE vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,107) = 91.18, **p < .000***	F (2,107) = 95.37, **p < .000***	F(2,107)= 29.2, **p < .000***
	50%		F (2,107) = 143, **p < .000***	F (2,107) = 67.7, **p < .000***
	75%		F (2,107) = 40, **p < .000***	F (2,107) = 85, **p < .000***
Sex	25%	F (1,107) = 1.81, p = .181	F (1,107) = 32.6, **p < .000***	F(1,107) = 8.7, **p = .004***
	50%		F (1,107) = 8.3, p = .443	F (1,107) = 4.2, **p = .042***
	75%		F (1,107) = .59, p = .399	F (1,107) = .01, p = .91
Language*sex	25%	F (2,107) = .594, p = .554	F (2,107) = 6.1, **p = .003***	F(2,107)= 2, p = .136
	50%		F (2,107) = 7.3, **p = .001***	F (2,107) = 1.8, p = .166
	75%		F (2,107) = 8, **p = .001***	F (2,107) = .359, p = .699

Furthermore, Table 7-9 shows that language had a significant effect on F1 and F2 differences between the three groups throughout the vowel. The Bonferroni post hoc tests revealed that throughout the vowel NE SQUARE was produced in a significantly lower position than both L2 and L1 productions (p < .000), which can be seen by comparing the three groups in Figure 7-16 below. As for F2 differences, the post hoc tests showed that L2 productions were similar to NE productions throughout the vowel: 25% (p = .306), 50% (p = 1), and 75% (p = .239). However, throughout the vowel, L1 productions were highly significantly fronter than both NE and L2 productions ($p < .000$). These findings suggest that the SA participants managed to perform a trajectory movement from a front mid vowel /ɛ/ towards a central vowel /ə/. However, since the participants' representations of these categories are in a higher position than those of NE (see chapter 4), L2 SQUARE productions were produced in a higher position. Thus, similarly to the previous vowel, the participants attempted to produce the diphthongal trajectory of this vowel, which they were able to perceive. However, their productions were NNL, which supports the prediction for the production of this vowel.

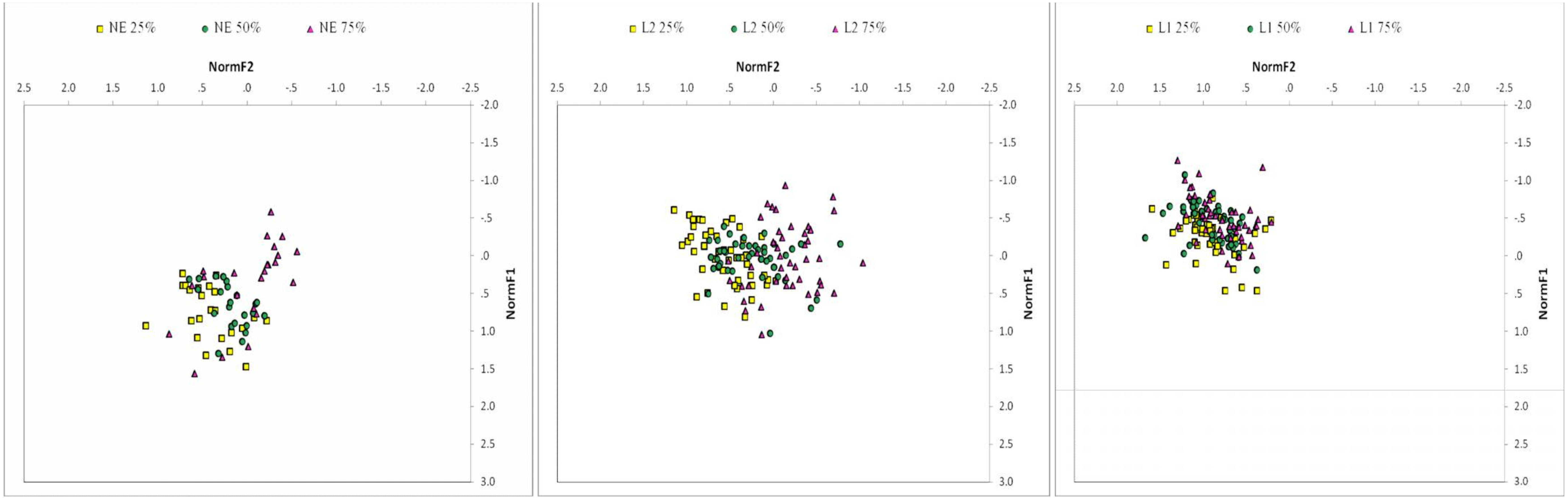

Figure 7-16: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English SQUARE vowel as produced by NE, L2, and L1 participants.

7.4.3.3 L2 (FACE, SQUARE)

Similar to the previous TC English contrasts, the SC (FACE, SQUARE) English contrast was predicted to be produced with distinct acoustic parameters DP by SA participants. In order to examine this prediction, the acoustic differences between L2 productions of these vowels were compared using a series of linear mixed models. Table 7-10 below presents a summary of the results, in which the two vowels highly significantly differ from each other in terms of quantity and quality.

Table 7-10: Summary of the results of the linear mixed model tests for L2 (FACE, SQUARE) vowels, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Vowel	25%	F (1,86) = 30.3, **p < .000***	F (1,86) = 18.8, **p < .000***	F(1,86)= 89.7, **p < .000***
	50%		F (1,86) = 197, **p < .000***	F (1,86) = 213, **p < .000***
	75%		F (1,86) = 105, **p < .000***	F (1,86) = 365, **p < .000***
Sex	25%	F (1,86) = 2.6, p = .106	F (1,86) = 3.9, **p = .049***	F(1,86) = .469, p = .495
	50%		F (1,86) = 4.15, **p = .045***	F (1,86) = 7.4, p = .39
	75%		F (1,86) = .817, p = .369	F (1,86) = .850, p = .359
Language*sex	25%	F (1,86) = .368, p = .546	F (1,86) = .013, p = .909	F(1,86)= 13.19, **p < .000***
	50%		F (1,86) = 7.3, **p = .001***	F (1,86) = 1.4, p = .232
	75%		F (1,86) = 11.6, **p = .001***	F (1,86) = .037, p = .849

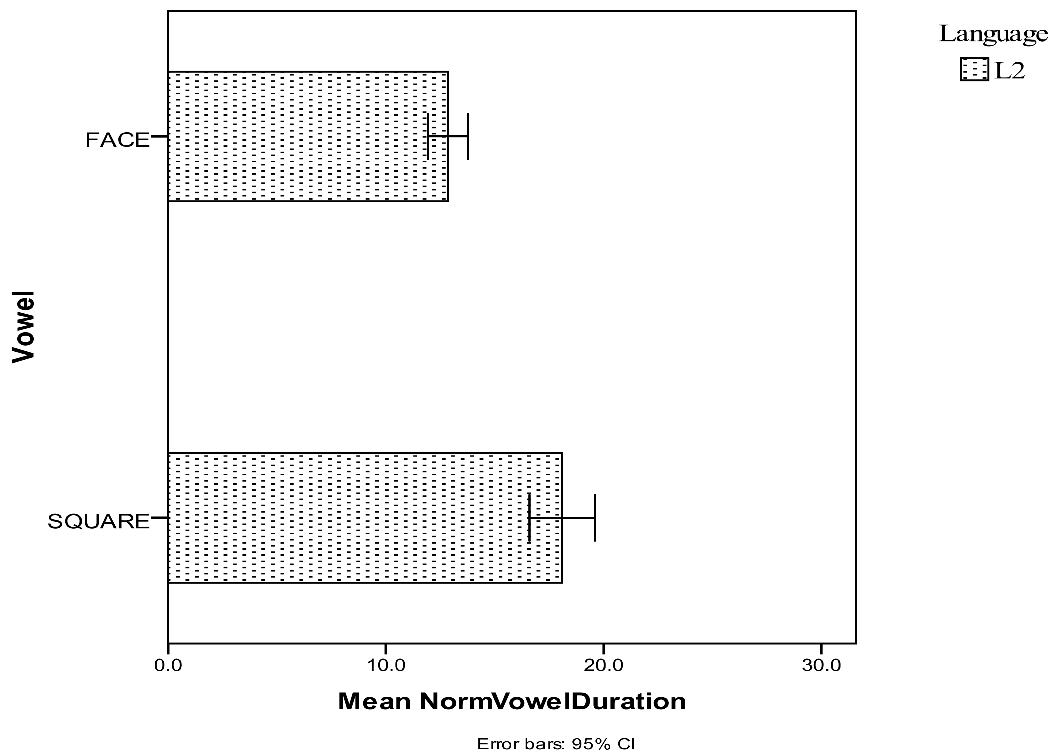

Figure 7-17: Mean normalised vowel duration of SA participants' L2 productions of English (FACE, SQUARE) contrast

Figure 7-17 above shows that the participants' L2 SQUARE productions are significantly longer in duration than their L2 FACE.

Figure 7-18 below presents the normalised F1 and F2 frequencies of 25%, 50%, and 75% of both vowels. As can be seen, L2 FACE and SQUARE productions are significantly different from each other throughout the vowel. In particular, FACE vowel productions project towards a high front quality, whereas the SQUARE productions project towards a mid central quality. Thus, those productions were diphthongal in nature showing a trajectory movement that is similar to that of NE productions. However, it can be argued that the amount of that movement and the on-glide and off-glide targets were adjusted to the participants' L2 vowel space, which is may be half way between that of NE and L1, particularly that L2 productions of the vowels examined in this study shared some acoustic properties with their L1 and others with NE. Therefore, the hypothesis for L2 production of this contrast is supported as the members of this contrast were produced with DP.

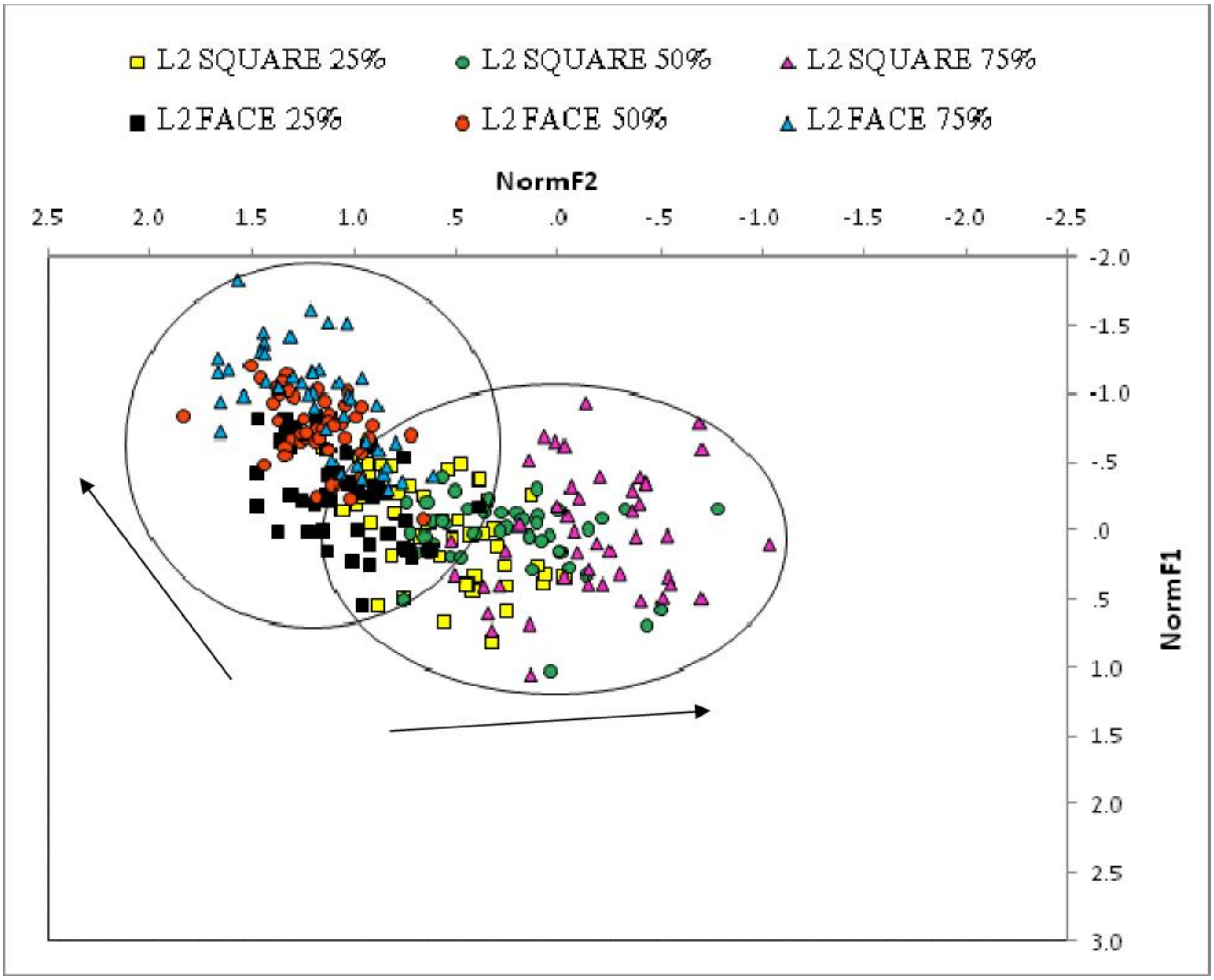

Figure 7-18: Mean normalised F1 and F2 values of SA participants' L2 productions of English (FACE, SQUARE) contrast.

7.4.4 The Production of the English (KIT, DRESS) contrast

7.4.4.1 KIT

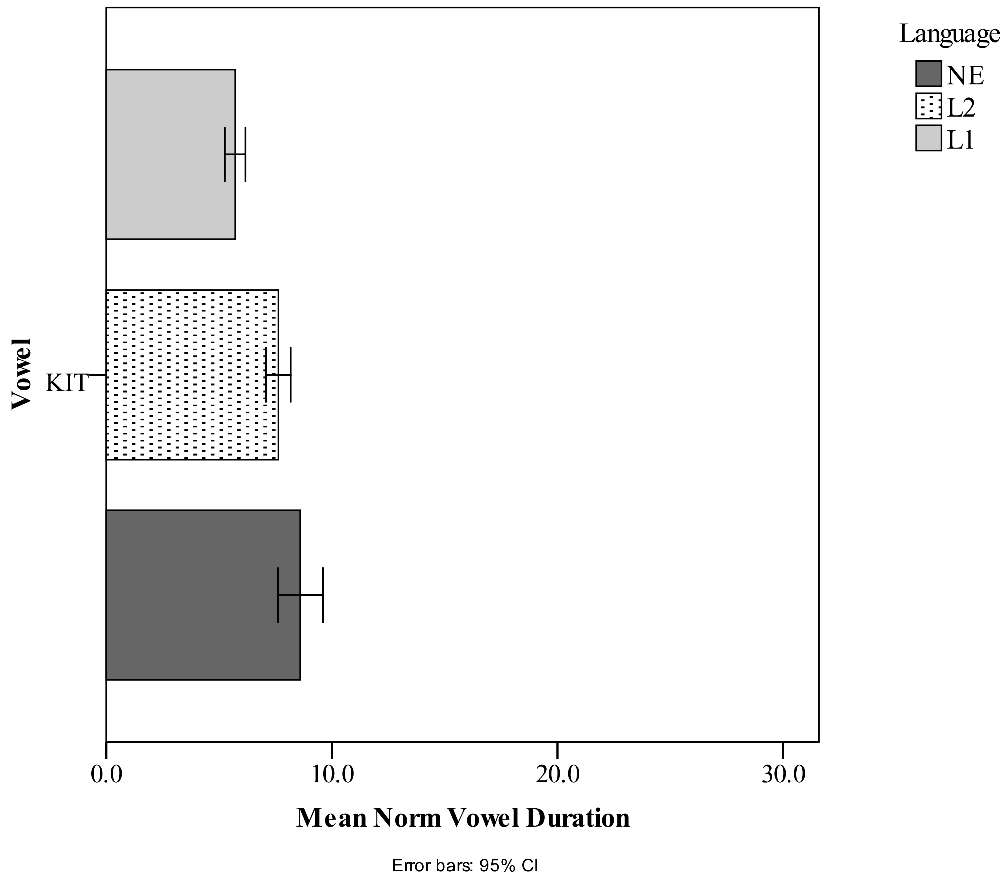

Figure 7-19: Mean normalised vowel duration of SA participants' L2 productions of English KIT vowel compared to NE productions and their L1/SA equivalent vowel /i/.

The production of the English KIT vowel was compared across the three language groups. The results in Table 7-11 below show that language had a significant effect on normalised vowel duration differences, in which L2 productions were slightly significantly shorter than NE productions (p = .038). Also, L1 productions of the /i/ vowel were significantly shorter than both NE and L2 productions (p < .000).

Figure 7-19 above shows that L2 KIT productions were produced with a duration that is half way between NE and L1 productions.

Table 7-11: Summary of the results of the linear mixed model tests for KIT vowel

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Language	F (2,105) = 20.9, **p < .000***	F (2,105) = 1.9, p = .141	F (2,105) = 10.97, **p < .000***
Sex	F (1,105) = 3.79, **p = .054***	F (1,105) = .203, p = .653	F (1,105) = 4.29, **p = .041***
Language*sex	F (2,105) = 1.14, p = .323	F (2,105) = 4.13, **p = .019***	F (2,105) = 3.64, **p = .03***

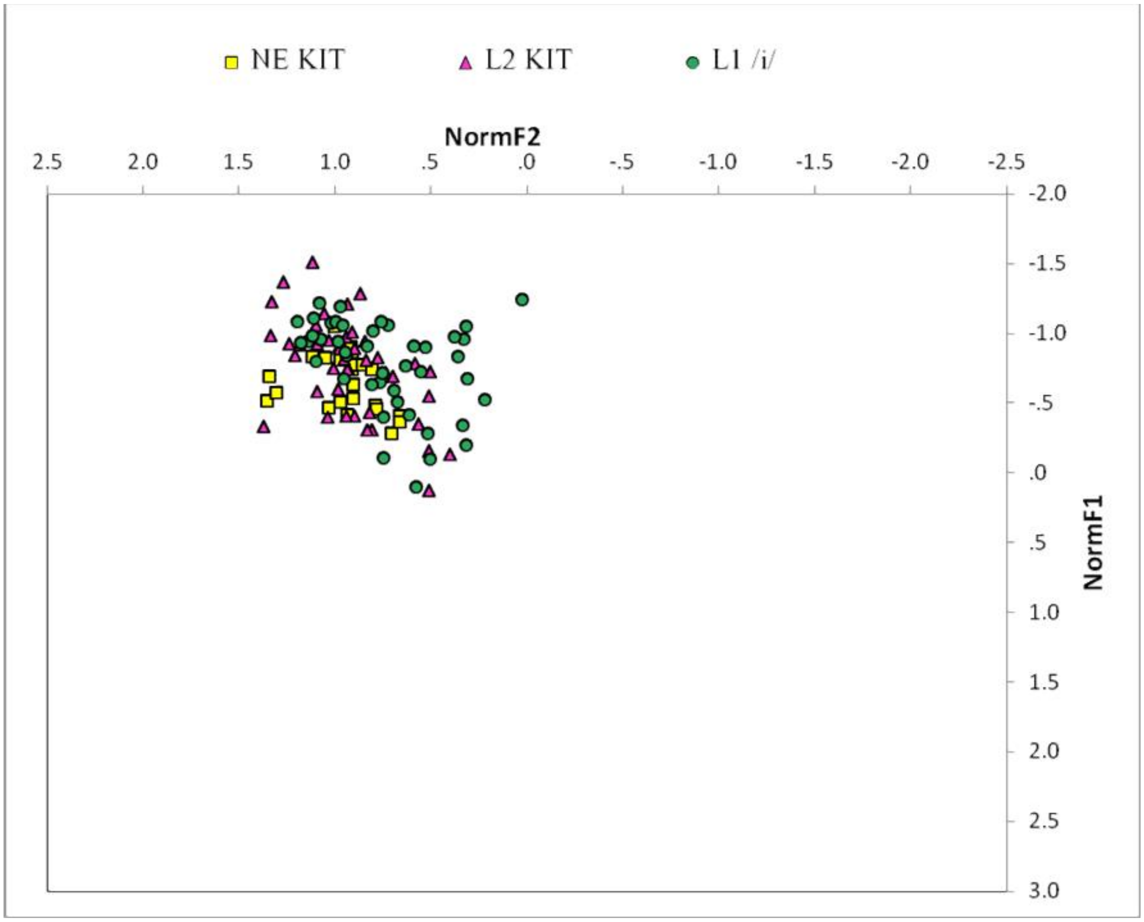

Figure 7-20: Mean normalised F1 and F2 values of SA participants' L2 productions of English KIT vowel compared to NE and L1 productions.

In terms of quality, Figure 7-20 above shows that there is a great overlap in the production of the three groups, particularly in height. Table 7-11 above shows that indeed language had a significant effect on F2 but not on F1 differences. Thus, the productions of the three groups were produced at a similar height. However, L1 productions were significantly more centralised than both NE ($p = .002$), and L2 productions ($p < .000$), which were similar ($p = 1$).

The results for L2 productions of the English KIT vowel suggest that SA participants' productions were qualitatively similar to NE productions, but quantitatively slightly shorter. This difference can be attributed to the fact that NE productions tend to have longer vowel durations when followed by voiced consonants, which was the case for the KIT vowel in the target word 'hid' (Arthur and House 1961). On the other hand, voicing of the adjacent consonants was found to have no effect on the duration of the vowel in Arabic or on L2 English productions of Arab learners (Mitleb 1982). Also, these findings support the PAM argument for the level of awareness of early L2 learners' perception, in which Best (2007: 23) argues that naïve listeners are not familiar with the phonetic details of the phonological distinctions of the L2, which is in this case the phonetic durational difference for vowels before voiced and voiceless consonants in English. Generally, L2 productions of KIT vowel were very close to NE productions, but they had NNL productions as they were slightly shorter, and this supports the hypothesis for this vowel.

7.4.4.2 DRESS

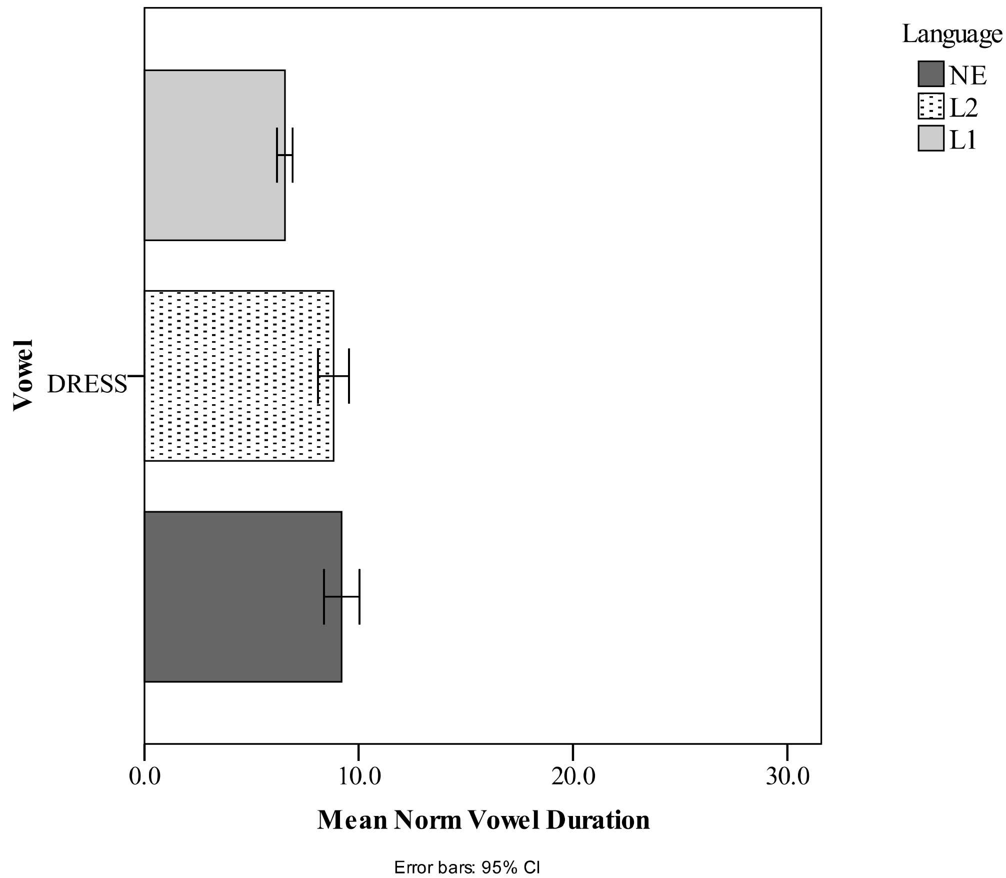

Figure 7-21: Mean normalised vowel duration of SA participants' L2 productions of English DRESS vowel compared to NE productions and their L1/SA equivalent vowel [e].

L2 productions of the English DRESS vowel were also compared to NE and L1 productions of the equivalent vowel [e], which was analysed as a phonetic variant of the SA /i/ vowel (see Chapter 4). The results for vowel duration are presented in Table 7-12 below, which shows that language had a significant effect on the durational differences. The Bonferroni post hoc test revealed that L2 productions were not significantly different from NE productions ($p = .631$), which can be seen in Figure 7-21 above. However, L1 productions of [e] vowel were significantly shorter than both NE and L2 productions ($p < .000$).

Table 7-12: Summary of the results of the linear mixed model tests for DRESS vowel

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Language	F (2,107) = 19.45, **p < .000***	F (2,107) = 40.9, **p < .000***	F (2,107) = 142, **p < .000***
Sex	F (1,107) = 4.4, **p = .038***	F (1,107) = .7.24, **p = .008***	F (1,107) = 1.9, p = .169
Language*sex	F (2,107) = .62, p = .54	F (2,107) = .32, p = .721	F (1,107) = 1.59, p = .208

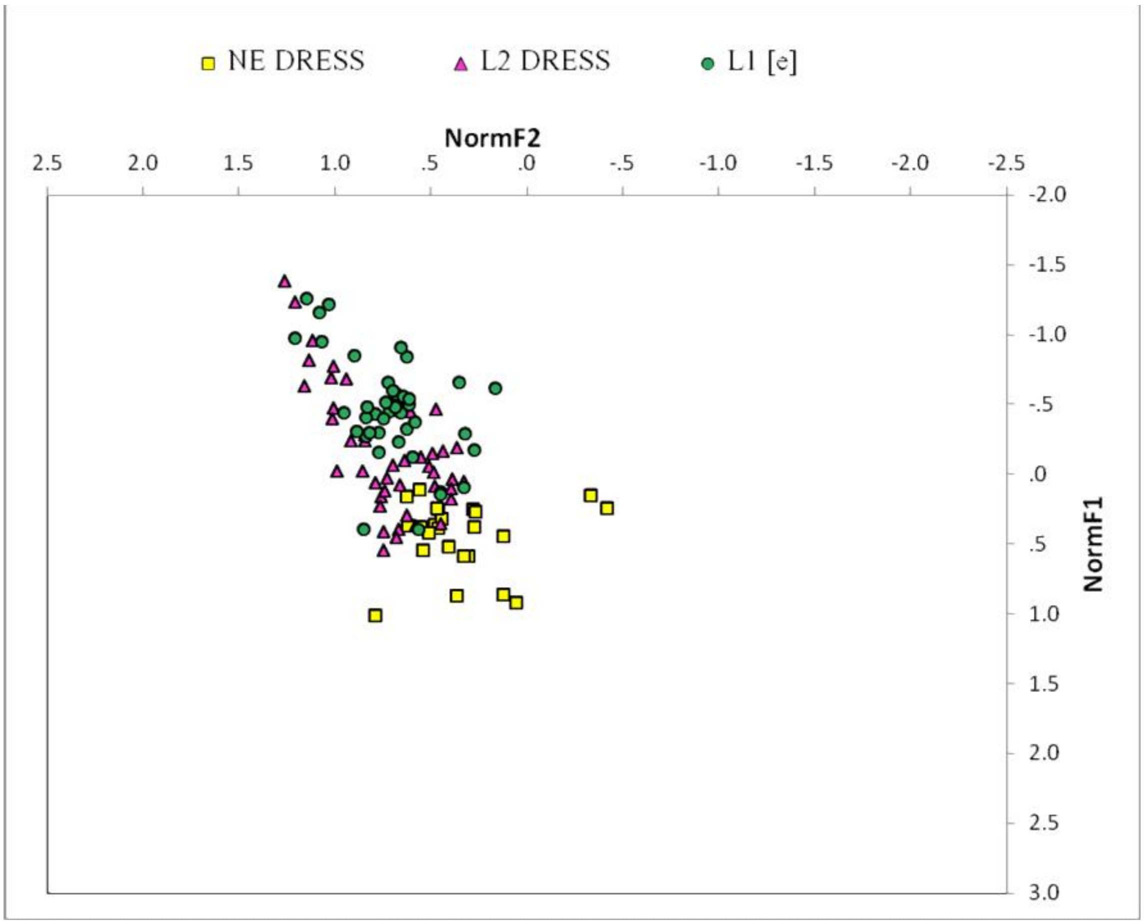

Figure 7-22: Mean normalised F1 and F2 values of SA participants' L2 productions of English DRESS vowel compared to NE and L1 productions.

As for quality, Figure 7-22 above shows that there is less overlap between the three groups than for the KIT vowel. The results in Table 7-12 above show that language had a significant effect on F1 and F2 differences. The Bonferroni post hoc tests showed that the participants' L2 productions were significantly higher and fronter ($p < .000$) than NE productions. Additionally, SA participants' L1 productions were slightly higher than their L2 productions ($p = .002$), but they were similar in frontness ($p = 1$).

These results showed that L2 DRESS productions were qualitatively significantly different from NE productions. Thus, similar to L2 KIT, this vowel was produced acoustically NNL, which supports the hypothesis for its predicted production. In the previous experiment, English DRESS received excellent identification close to that of NE. This was suggested to be the result of acoustic similarity between NE DRESS and the SA phonological category /i/, rather than its phonetic category [e]. A closer examination of Figure 7-21 in the previous section and Figure 7-22 above reveals that NE DRESS productions and L1 /i/ productions do not overlap in the vowel space. Thus, the high correct identification of DRESS vowel cannot be attributed to its similarity to SA /i/ or [e] vowels. Additionally, the present study showed that the English KIT vowel, which was analysed as the closer to the L1 category in the PAT, was produced closer to the L1 category /i/ than English DRESS. Thus, the question of why English DRESS was highly accurately identified more than KIT remains open for further discussion in section 7.5.3 below.

7.4.4.3 L2 (KIT, DRESS)

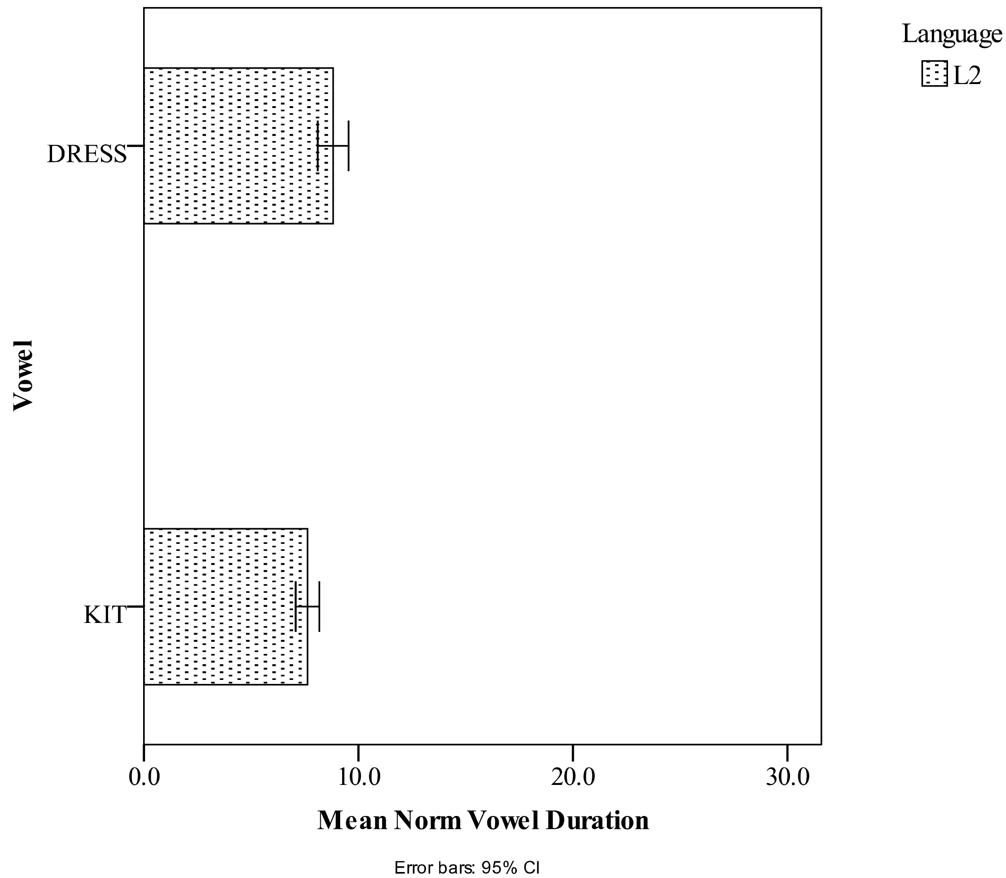

Figure 7-23: Mean normalised vowel duration of SA participants' L2 productions of English (KIT, DRESS) contrast

The hypothesis for the L2 production of the CG (KIT, DRESS) contrast was that the members of this contrast will be produced with overlapping OP. Similar to the previous contrasts, a series of linear mixed models were run to compare the two vowels. First, Table 7-13 below shows that L2 KIT vowel was produced with a significantly shorter duration than L2 DRESS vowel, which can be seen in Figure 7-23 above. Also, the table shows that L2 productions of these vowels were significantly different in height and frontness. Figure 7-24 below shows that indeed L2 KIT productions are significantly higher and fronter than L2 DRESS productions. Thus, the hypothesis for this contrast was not supported as the two vowels were acoustically distinct from each other DP.

Table 7-13: Summary of the results of the linear mixed model tests for L2 (KIT, DRESS) productions

Fixed effects	Dependent variables		
	Norm vowel duration	Norm F1	Norm F2
Vowel	F (1,85) = 6.8, **p < .011***	F (1,85) = 57, **p < .000***	F (1,85) = 21.5, **p < .000***
Sex	F (1,85) = 9.2, **p = .003***	F (1,85) = .013, p = .908	F (1,85) = 4.31, **p = .041***
Language*sex	F (1,85) = .004, p = .95	F (1,85) = 6.31, **p = .014***	F (1,85) = 8.82, **p = .004***

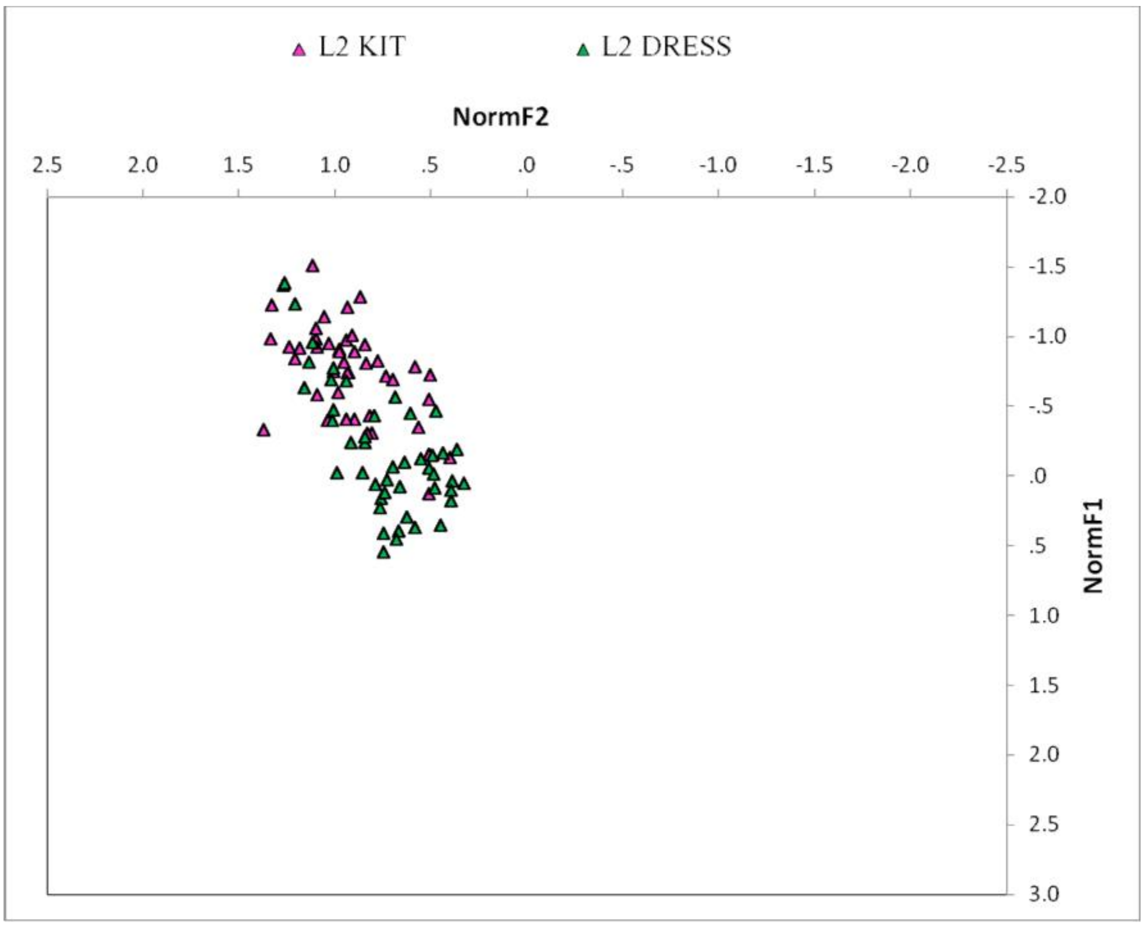

Figure 7-24: Mean normalised F1 and F2 values of SA participants' L2 productions of English (KIT, DRESS) contrast

Although L2 productions of English (KIT, DRESS) vowels were NNL, the relative differences between these vowels were produced the same by NE and L2 participants. This means that NE DRESS productions were generally longer, lower, and more centralised than NE KIT productions. Similarly, L2 productions of English DRESS were longer, lower, and more centralised. Thus, it can be argued that SA participants did realise and produce the relative acoustic differences between these vowels similar to NE participants. However, L2 productions were adjusted to the participants' phonological space, which was in most cases half way between that of their L1 and NE participants.

7.4.5 The Production of the English (GOAT, THOUGHT) contrast

7.4.5.1 GOAT

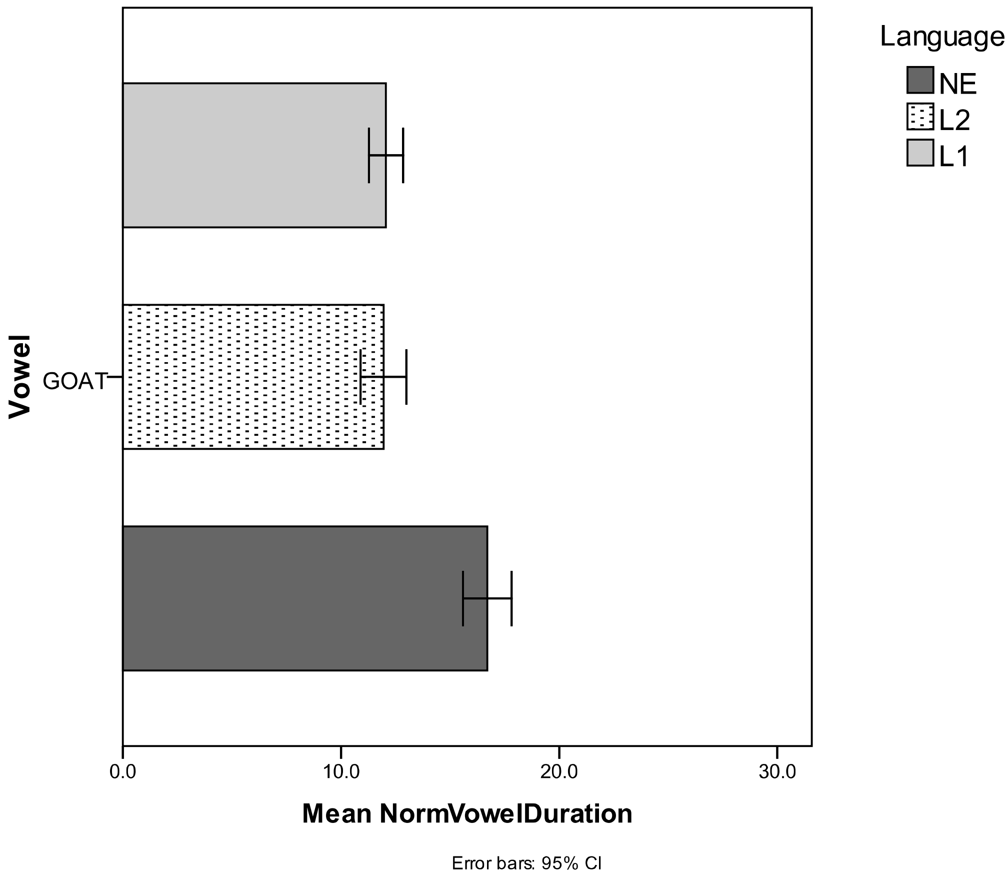

Figure 7-25: Mean normalised vowel duration of SA participants' L2 productions of English GOAT vowel compared to NE productions and their L1/SA equivalent vowel /oː/.

Figure 7-25 above and Table 7-14 below show that language had a significant effect on the durational differences of English GOAT productions compared to its SA equivalent mid long vowel /oː/. The Bonferroni post hoc test showed that NE productions were significantly longer than both the L2 and L1 productions ($p < .000$), which were produced with a similar vowel duration ($p = 1$). The great durational match between the L2 GOAT productions and the L1 /oː/ vowel suggests that L2 productions were monophthongal (if we assume that monophthongs are shorter than diphthongs), which would explain production of L2 GOAT with a shorter duration than NE productions.

Table 7-14: Summary of the results of the linear mixed model tests for GOAT vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,108) = 25.5, **p < .000***	F (2,108) = 2, p = .131	F(2,108)= 315.7, **p < .000***
	50%		F (2,108) = 14.9, **p < .000***	F (2,108) = 225, **p < .000***
	75%		F (2,108) = 27.7, **p < .000***	F (2,108) = 51, **p < .000***
Sex	25%	F (1,108) = 2.42, p = .122	F (1,108) = 13.5, **p < .000***	F(1,108) = .014, p = .905
	50%		F (1,108) = 1.08, p = .3	F (1,108) = .006, p = .94
	75%		F (1,108) = 1.48, p = .225	F (1,108) = .058, p = .81
Language*sex	25%	F (2,108) = .7, p = .499	F (2,108) = 2.65, p = .075	F(2,108)= 4.87, **p = .009***
	50%		F (2,108) = 5.27, **p = .007***	F (2,108) = .59, p = .556
	75%		F (2,108) = .616, p = .542	F (2,108) = 2.42, p = .093

Figure 7-26 below presents the normalised F1 and F2 frequencies of NE, L2, and L1 productions. Similar to the diphthongal (FACE, SQUARE) vowels, the measurements were extracted from three points throughout the vowel to show its trajectory movement. As can be seen, NE productions exhibit a trajectory movement from a mid central vowel quality /ə/ into a mid-high central vowel /ʊ/. On the other hand, L2 GOAT productions exhibit the small trajectory movement which would be expected for a monophthongal production. Unexpectedly, L1 /oː/ productions exhibit a small trajectory movement which contradicts what would be expected for Arabic long vowels. Table 7-14 above shows that language had a significant effect on F1 differences of midpoint productions and towards the end of the vowel. The Bonferroni post hoc tests showed that L2 GOAT productions were significantly lower than NE productions ($p < .000$), but not significantly different from their L1 productions across the three points: ($p = 1$), ($p = .301$), and ($p = .073$), respectively. Also, language had a significant effect on F2 differences throughout the vowel, in which L2 GOAT productions were significantly backer than NE productions across the three points ($p < .000$). Additionally, L2 productions were significantly fronter than their L1 productions at the beginning ($p < .000$) and midpoint ($p = .003$) of the vowel. However, towards the end of the vowel, L2 and L1 productions were produced in a similar position on the F2 dimension ($p = 1$).

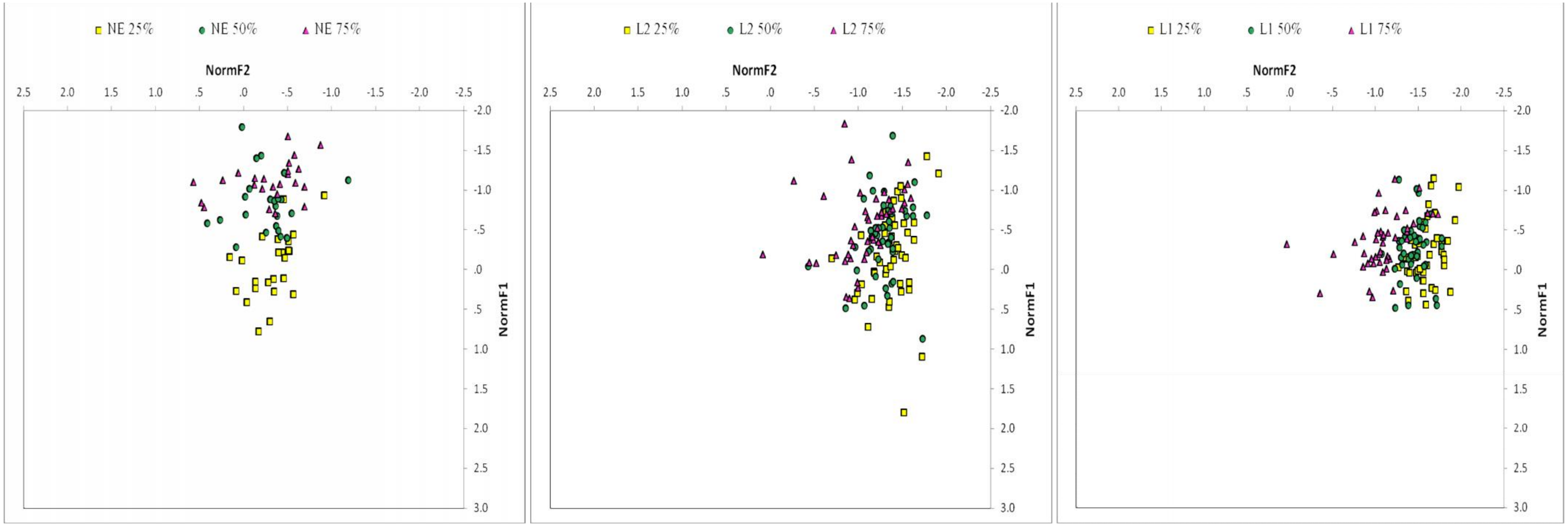

Figure 7-26: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English GOAT vowel as produced by NE, L2, and L1 participants.

Generally, the above results showed that L2 English GOAT productions were NNL as they were significantly shorter, lower, and backer than NE productions, which supports the hypothesis for the production of this vowel. As predicted, since GOAT was not given a high goodness of fit rating to its L1 equivalent vowel /oː/, L2 productions of this vowel shared some acoustic properties with the L1 vowel, i.e. duration, height, and the place of the end of the vowel. The difficulty in producing native-like GOAT productions can be attributed to the effect of the L1 vowel system, which does not have a phonological schwa vowel. In chapter 4, the SA schwa category was analysed as a phonetic variant of the short vowel /i/, where it was suggested to surface as a phonetic variant of other SA short vowels. The participants attempted to produce a more centralised on-glide than that of their L1 productions. However, towards the end of their L2 productions the participants tended towards their L1 phonological category /oː/. In sum, it was more difficult for the participants to identify and discriminate this vowel in the previous chapter than other English vowels, as well as to produce it.

7.4.5.2 THOUGHT

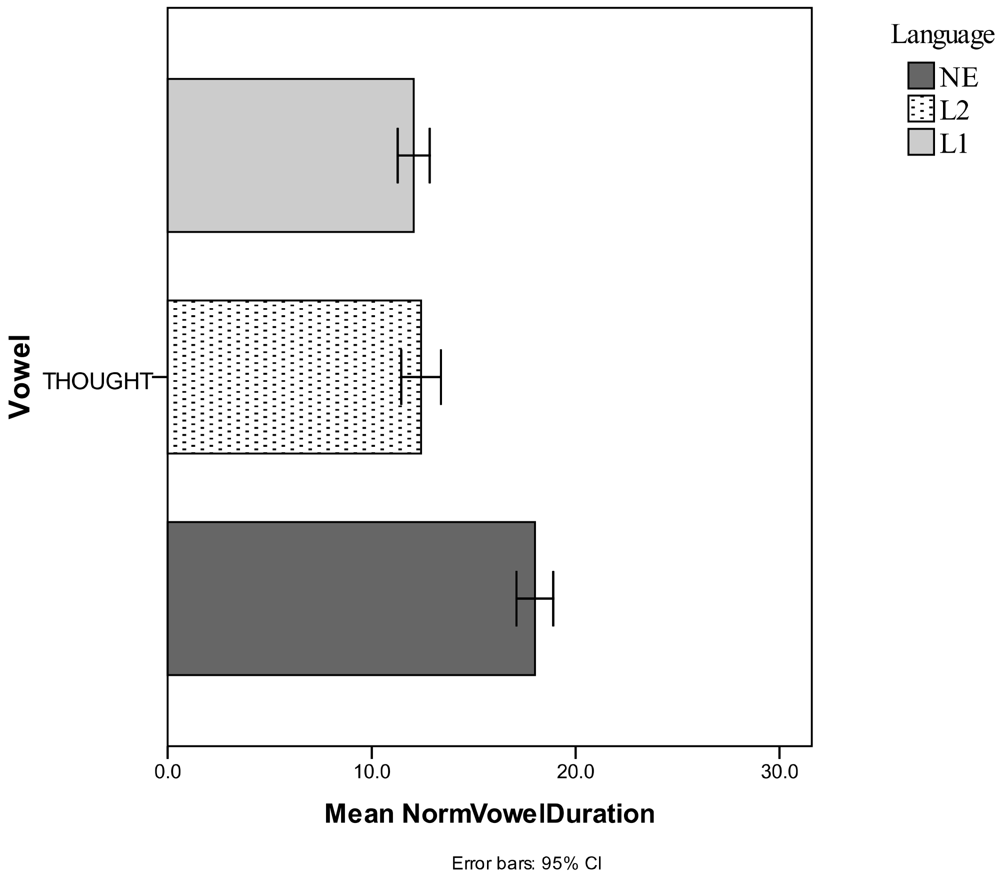

Figure 7-27: Mean normalised vowel duration of SA participants' L2 productions of English THOUGHT vowel compared to NE productions and their L1/SA equivalent vowel /o:/.

Figure 7-27 above presents mean normalised vowel duration of the English THOUGHT vowel. As can be seen, the durational differences between the three language groups show a similar pattern to that of English GOAT vowel, in which NE productions were produced longer than L2, and L1 productions of /o:/ vowel. Table 7-15 below shows that language had a significant effect on these differences, in which NE productions were significantly longer than both L2 and L1 productions ($p < .000$). Similar to the previous vowel, L2 THOUGHT and L1 /o:/ vowels were produced with a similar vowel duration ($p = 1$). Thus, L2 (GOAT, THOUGHT) vowels were produced quantitatively with a similar vowel duration that was the same as their L1 /o:/ category.

Table 7-15: Summary of the results of the linear mixed model tests for THOUGHT vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,107) = 47.3, **p < .000***	F (2,107) = 14.4, **p < .000***	F(2,107)= 29, **p < .000***
	50%		F (2,107) = 5.9, **p = .003***	F (2,107) = 33, **p < .000***
	75%		F (2,107) = 3.86, **p = .024***	F (2,107) = 13.6, **p < .000***
Sex	25%	F (1,107) = 3.22, p = .076	F (1,107) = 10.2, **p = .002***	F(1,107) = 1.52, p = .699
	50%		F (1,107) = 5.7, **p = .019***	F (1,107) = .15, **p = .042***
	75%		F (1,107) = .13, p = .719	F (1,107) = .01, p = .911
Language*sex	25%	F (2,107) = 2.7, p = .071	F (2,107) = 2.94, **p = .057***	F(2,107)= 1.84, p = .163
	50%		F (2,107) = .31, p = .729	F (2,107) = 2.5, p = .083
	75%		F (2,107) = 1.9, p = .152	F (2,107) = 3.49, **p = .034***

In terms of quality, English THOUGHT vowel and its SA equivalent vowel /oː/ are monophthongs. However, the qualitative analysis of these productions was similar to the previous diphthong GOAT, i.e. frequency measurements presented at three points, in order to have a comparable analysis for the production of the two members of English (GOAT, THOUGHT) contrast. Similarly, Figure 7-28 below presents the frequency measurements of the three groups, which were extracted from three points across the vowel. Unexpectedly, L1 /oː/ productions exhibit a greater trajectory movement than NE THOUGHT and this contradicts the expected lack of diphthongisation for Arabic long vowels compared to NE. The results in Table 7-15 show that language had a significant effect on F1 and F2 differences across the three points. As would be expected for NE monophthongal productions, the measurements of the three points cluster in a mid-high back position. Compared to NE productions, L2 THOUGHT productions were significantly lower at the beginning and towards the midpoint of the vowel ($p < .000$), ($p = .024$), but similar towards the end ($p = .222$). Additionally, L2 productions were fronter than NE productions throughout the vowel ($p < .000$). On the other hand, L2 productions had a similar height to their L1 productions throughout the vowel: ($p = .602$), ($p = 1$), and ($p = .796$), respectively. However, L2 productions were fronter than their L1 equivalent productions on the F2 dimension towards the midpoint ($p < .000$), but similar towards the end ($p = 1$).

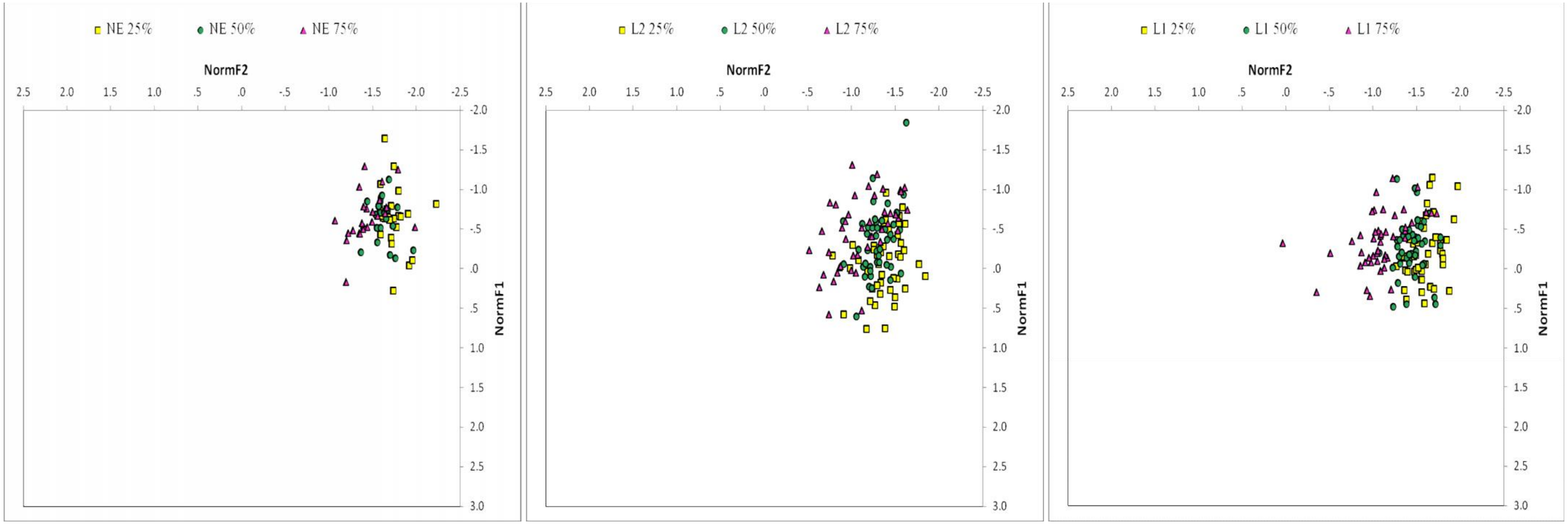

Figure 7-28: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English THOUGHT vowel as produced by NE, L2, and L1 participants.

The results for this vowel show that the productions of the three groups were monophthongal in nature. However, L2 productions were NNL as they were significantly shorter, lower, and fronter than NE productions, which supports the hypothesis for this vowel. Unlike what was predicted, L2 productions of the English THOUGHT vowel were not closer to the L1 category /o:/ than English GOAT since both were fronter than the L1 vowel. The similarity in the L2 production of the members of the English (GOAT, THOUGHT) contrast compared to NE and L1 productions suggest that both L2 productions overlap acoustically, which will be examined in the following section.

7.4.5.3 L2 (GOAT, THOUGHT)

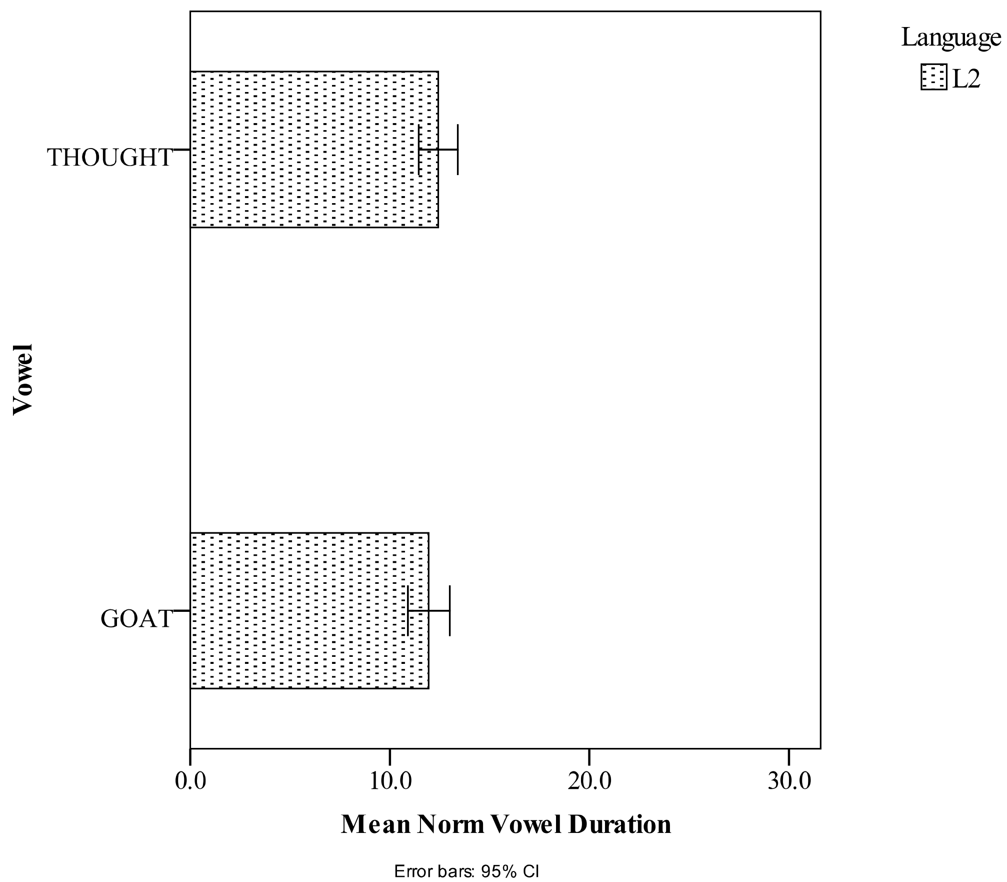

Figure 7-29: Mean normalised vowel duration of SA participants' L2 productions of English (GOAT, THOUGHT) contrast

Similar to the previous CG contrast, L2 productions of English (GOAT, THOUGHT) vowels were predicted to be produced with overlapping acoustic parameters OP. Figure 7-29 above and Figure 7-30 below show that both vowels are quantitatively and qualitatively similar. The results of the linear mixed model tests comparing the two vowels in Table 7-16 below show that there are no significant differences in the L2 productions of these vowels. Thus, the hypothesis for the L2 production of this contrast is supported as these vowels were produced as acoustically similar.

Table 7-16: Summary of the results of the linear mixed model tests for L2 (GOAT, THOUGHT) vowels, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Vowel	25%	F (1,85) = .21, p = .642	F (1,85) = 1.4, p = .226	F(1,85)= .002, p = .962
	50%		F (1,85) = .849, p = .359	F (1,85) = .218, p = .642
	75%		F (1,85) = .939, p = .335	F (1,85) = .343, p = .56
Sex	25%	F (1,85) = 5.9, **p = .017***	F (1,85) = 28.6, **p < .000***	F(1,85) = 8.99, **p = .004***
	50%		F (1,85) = .518, p = .474	F (1,85) = .173, p = .678
	75%		F (1,85) = 1.57, p = .213	F (1,85) = 3.61, p = .061
Language*sex	25%	F (1,85) = .364, p = .548	F (1,85) = .585, p = .446	F(1,85)= .004, p = .953
	50%		F (1,85) = 8.29, **p = .005***	F (1,85) = .032, p = .858
	75%		F (1,85) = .03, p = .863	F (1,85) = .016, p = .901

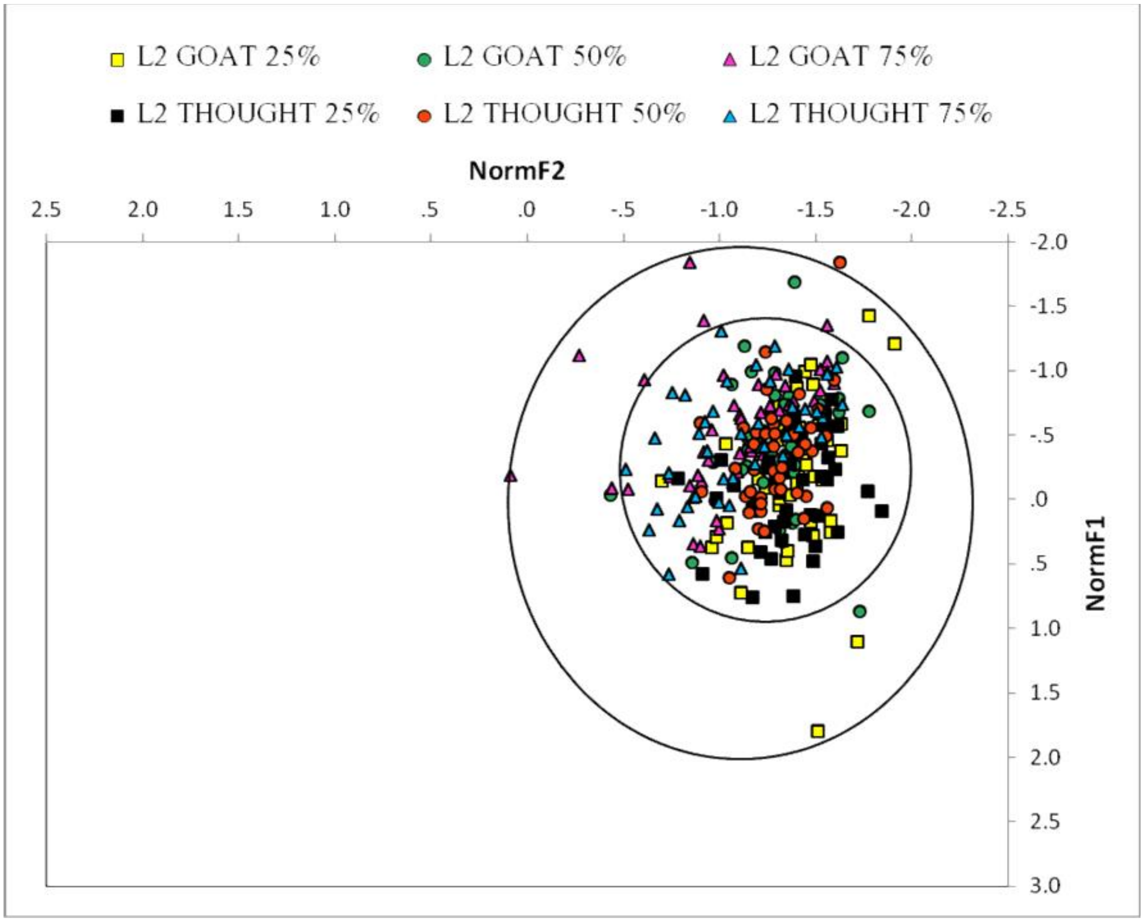

Figure 7-30: Mean normalised F1 and F2 values of SA participants' L2 productions of English (GOAT, THOUGHT) contrast.

7.4.6 The Production of the English (FLEECE, NEAR) contrast

7.4.6.1 FLEECE

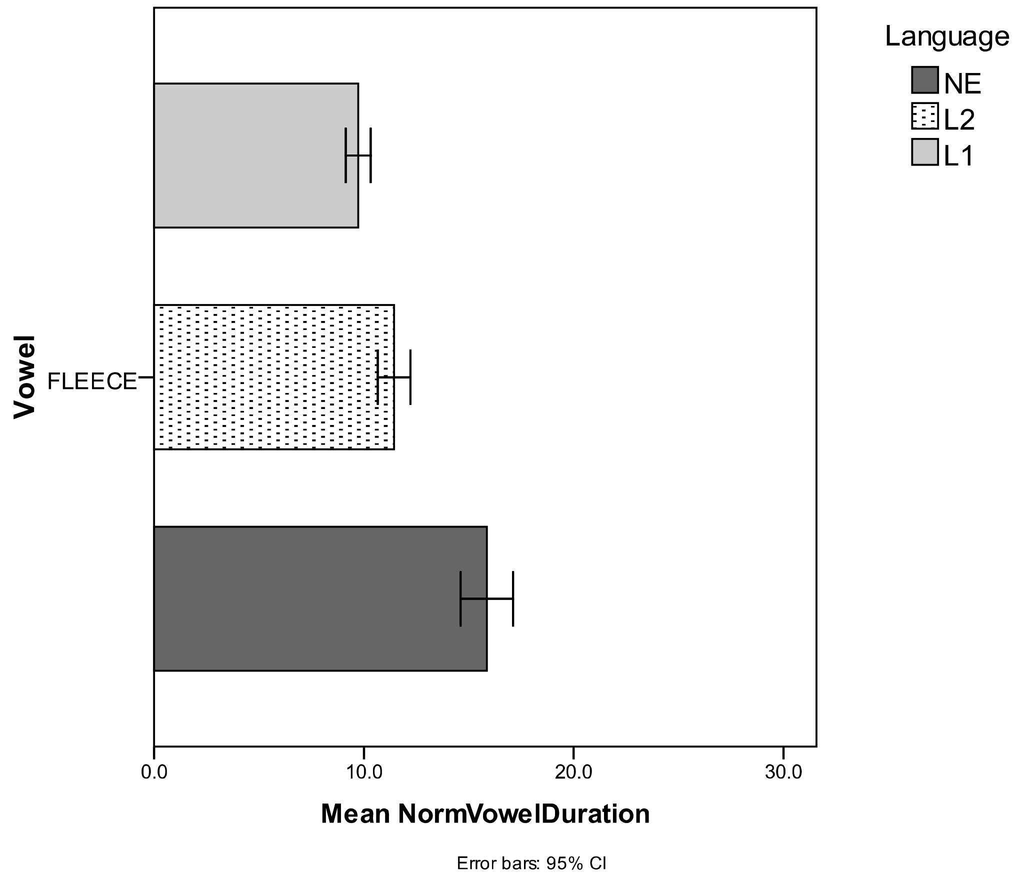

Figure 7-31: Mean normalised vowel duration of SA participants' L2 productions of English FLEECE vowel compared to NE productions and their L1/SA equivalent vowel /iː/.

Figure 7-31 above presents the mean of the normalised vowel duration of English FLEECE, comparing NE, L2 and L1 equivalent productions of the SA /iː/ vowel. The durational pattern for this vowel is similar to the previous vowels, in which NE productions tend to be longer than L2 and L1 productions. Table 7-17 below shows that Language had a significant effect on the durational differences, where L2 productions were significantly shorter than NE productions ($p < .000$) but longer than their L1 productions ($p = .007$).

Table 7-17: Summary of the results of the linear mixed model tests for FLEECE vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,108) = 47.3, **p < .000***	F (2,108) = 2.01, p = .138	F(2,108) = 1.77, p = .174
	50%		F (2,108) = 2.87, p = .061	F (2,108) = 5.42, **p = .006***
	75%		F (2,108) = 6.48, **p = .002***	F (2,108) = 2.33, p = .102
Sex	25%	F (1,108) = .34, p = .556	F (1,108) = 1.27, p = .261	F(1,108) = .572, p = .451
	50%		F (1,108) = 3.95, **p = .049***	F (1,108) = .001, p = .972
	75%		F (1,108) = 11.5, **p = .001***	F (1,108) = 2.29, p = .133
Language*sex	25%	F (2,108) = .035, p = .966	F (2,108) = 5.88, **p = .004***	F(2,108)= 1.39, p = .251
	50%		F (2,108) = 7.44, **p = .001***	F (2,108) = 1.08, p = .343
	75%		F (2,108) = 7.55, **p = .001***	F (2,108) = .781, p = .46

As for quality, the analysis of the FLEECE productions was similar to the analysis of the previous diphthongal vowels, in order to compare the production of this vowel to its diphthongal contrast NEAR. Table 7-17 above shows that language did not have any significant effect on F1 differences at the beginning and towards the midpoint of the vowel. However, language effect was significant towards the end of the vowel, in which L2 productions were significantly lower than NE productions (p = .002) but similar in height to their L1 equivalent productions (p = 1), as can be seen in Figure 7-32 below.

The table also shows that the language*sex interaction effect on F1 differences was significant at the end point of the vowel. Further examination showed that L2 females' productions of English FLEECE were lower than L2 males' and NE productions. Thus, the female SA participants were responsible for the height difference between L2 and NE productions. Figure 7-32 below shows two outliers, which were checked and found to belong to the female SA S5 participant, and her productions affected the height results. In terms of frontness, language effect was only significant at the midpoint of the vowel, in which L2 productions were slightly significantly backer than NE productions ($p = .044$), but similar to their L1 productions ($p = 1$). Similarly to SA /oː/, the productions of SA /iː/ vowel show a small diphthongisation, which is greater than that of NE FLEECE productions. Diphthongisation in Arabic vowels was not examined in any previous research, and the findings of this study shows interesting patterns worth of further analysis in the future.

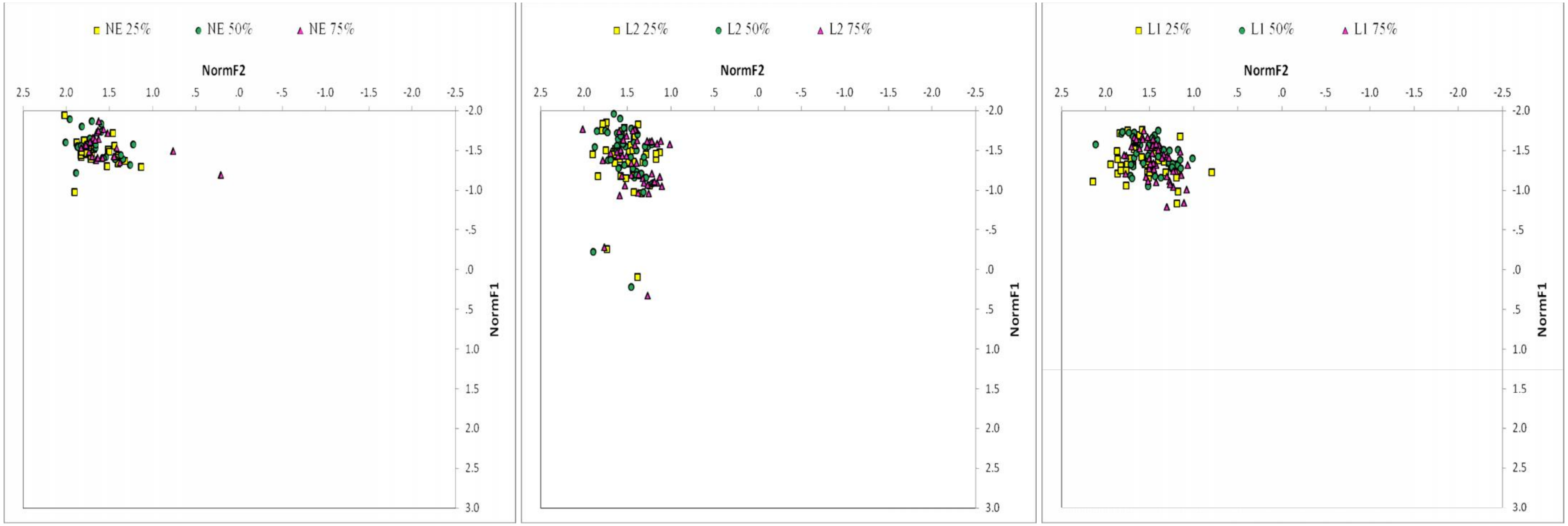

Figure 7-32: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English FLEECE vowel as produced by NE, L2, and L1 participants.

Generally, the production of the FLEECE vowel was monophthongal in nature for the three language groups. However, L2 and L1 productions were significantly shorter and backer than NE productions. Therefore, L2 productions were NNL, which contradicts the prediction for this vowel. This contradiction can be attributed to the acoustic differences between NE and L1 productions, which yielded L2 productions closer to the L1 category since English FLEECE vowel had a high similarity to that L1 vowel.

7.4.6.2 NEAR

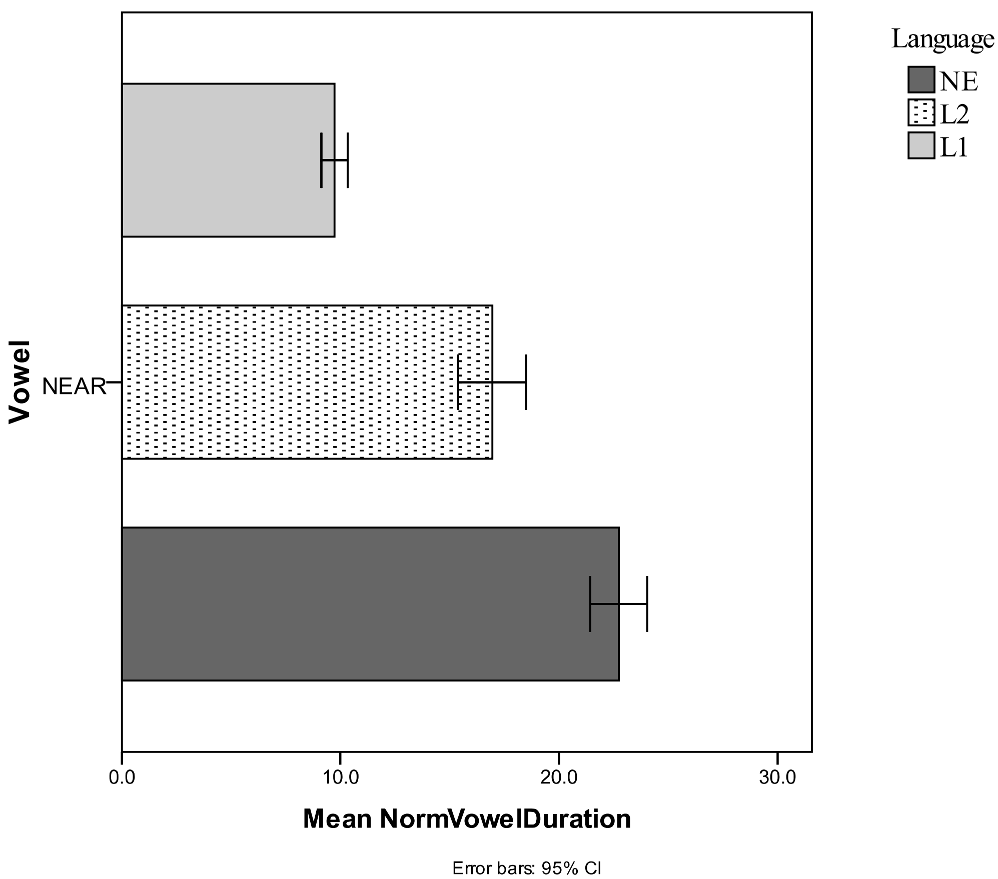

Figure 7-33: Mean normalised vowel duration of SA participants' L2 productions of English NEAR vowel compared to NE productions and their L1/SA equivalent vowel /iː/.

The mean normalised duration of NEAR diphthongal vowel productions in Figure 7-33 above shows that L2 productions are longer than their L1 /iː/ productions and shorter than NE productions. Table 7-18 below shows that language had a significant effect on these differences, in which the three groups differed significantly from each other ($p < .000$).

Table 7-18: Summary of the results of the linear mixed model tests for NEAR vowel, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Language	25%	F (2,108) = 95.5, **p < .000***	F (2,108) = 23, **p < .000***	F(2,108) = 24, **p < .000***
	50%		F (2,108) = 108, **p < .000***	F (2,108) = 89.6, **p < .000***
	75%		F (2,108) = 116, **p < .000***	F (2,108) = 248, **p < .000***
Sex	25%	F (1,108) = 2.06, p = .154	F (1,108) = .42, p = .517	F(1,108) = .114, p = .736
	50%		F (1,108) = .107, p = .745	F (1,108) = 1.82, p = .18
	75%		F (1,108) = .002, p = .962	F (1,108) = .715, p = .4
Language*sex	25%	F (2,108) = 1.29, p = .278	F (2,108) = 1.68, p = .191	F(2,108)= 1.53, p = .221
	50%		F (2,108) = 5.27, **p = .007***	F (2,108) = .296, p = .745
	75%		F (2,108) = 10.8, **p < .000***	F (2,108) = .074, p = .928

As for quality, the table also shows that language effect on F1 and F2 differences was significant throughout the vowel. A comparison between the three groups in Figure 7-34 below shows that L2 productions exhibit a very similar trajectory movement to the one produced by NE participants, i.e. a projection from a high front vowel /ɪ/ into a mid central schwa vowel. The Bonferroni post hoc tests revealed that indeed L2 productions were similar to NE productions on the F2 dimension throughout the vowel: ($p = 1$). However, L2 productions were significantly higher than those of NE: ($p = .004$), ($p < .000$), and ($p < .000$), respectively. On the other hand, L2 productions of the NEAR vowel were highly significantly lower and fronter than their L1 productions throughout the vowel ($p < .000$). Thus, L2 productions were acoustically significantly different from their L1 productions of the /iː/ vowel.

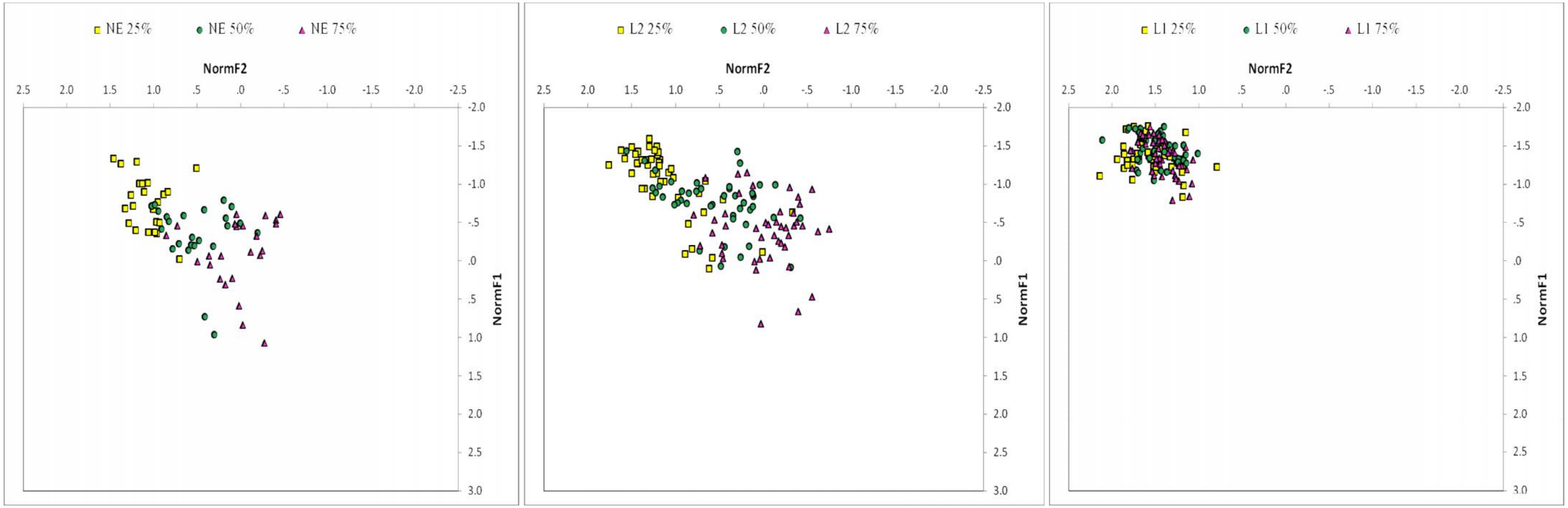

Figure 7-34: Mean normalised F1 and F2 frequencies of 25%, 50%, and 75% of English NEAR vowel as produced by NE, L2, and L1 participants.

The results of the present study showed that the uncategorised English vowel NEAR was produced by SA participants with a similar diphthongal projection, but on a higher position on the F1 dimension. Similarly to the English SQUARE diphthong, the participants have higher representations for the on-glide and off-glide components of NEAR vowel. Therefore, they were able to produce the trajectory movement in NEAR and SQUARE, which they were argued to be able to perceive. L2 productions were also significantly shorter than NE productions. Thus, L2 productions were NNL, which supports the hypothesis for this vowel. Unlike what was predicted however, those L2 productions exhibited a diphthongal trajectory which was similar to NE productions. The implications of the latter finding will be discussed further in section 7.5 below.

7.4.6.3 L2 (FLEECE, NEAR)

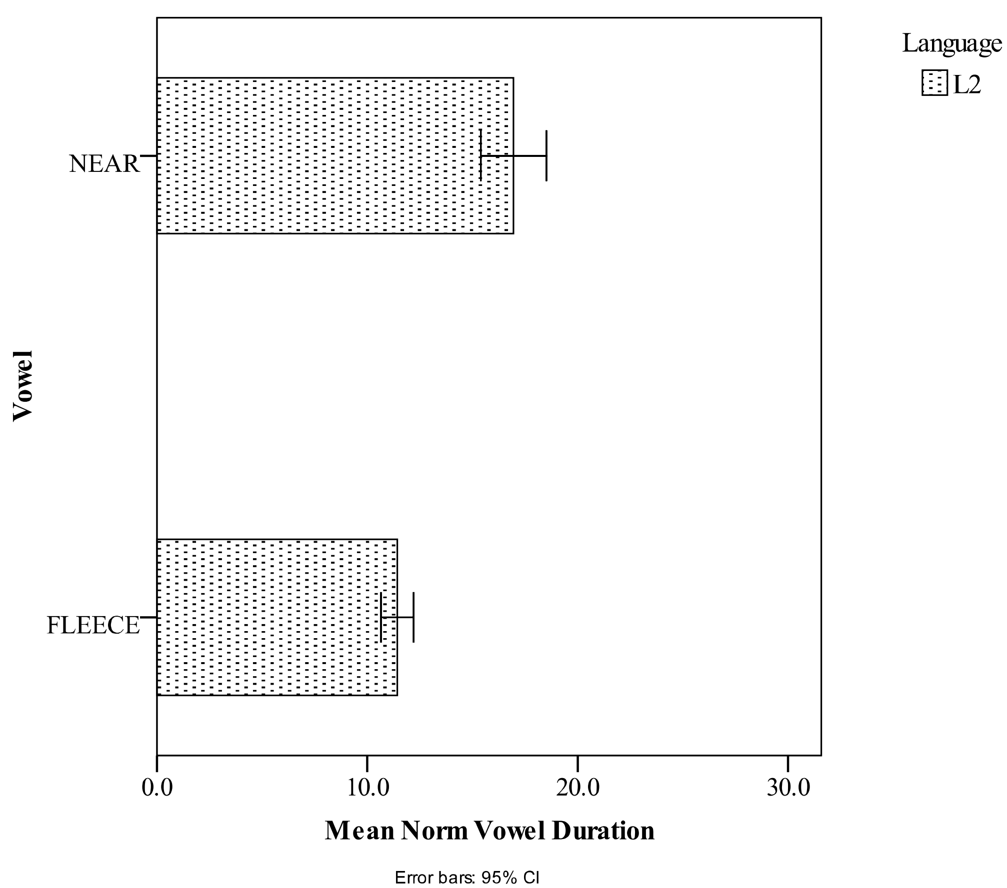

Figure 7-35: Mean normalised vowel duration of SA participants' L2 productions of English (FLEECE, NEAR) contrast

The hypothesis for the L2 production of the UC English (FLEECE, NEAR) contrast was that the members of this contrast will be produced with overlapping acoustic parameters OP since both vowels were categorised as a single SA vowel /iː/, with FLEECE being the categorised vowel and NEAR the uncategorised one. Figure 7-35 above compares L2 productions of English (FLEECE, NEAR) contrast in terms of duration. As can be seen, the participants' L2 FLEECE productions are shorter than their productions of NEAR vowel. Also, Figure 7-36 below shows that L2 productions of the two vowels differ from each other in terms of F1 and F2. These differences are reflected in the results of the linear mixed model tests in Table 7-19 below, in which the L2 FLEECE vowel was significantly shorter, higher, and fronter than L2 NEAR productions throughout the vowel. Thus, the results of the present study do not support the hypothesis for this contrast. As shown, L2 productions of these vowels were produced with distinct parameters DP, where FLEECE vowel was monophthongal in nature whereas NEAR was diphthongal.

Table 7-19: Summary of the results of the linear mixed model tests for L2 (FLEECE, NEAR) vowels, this includes results for tests which were run for the frequency measurements of 25%, 50%, and 75% of the vowel

Fixed effects	part of the vowel	Dependent variables		
		Norm vowel duration	Norm F1	Norm F2
Vowel	25%	F (1,86) = 31.9, **p < .000***	F (1,86) = 10.6, **p = .002***	F(1,86)= 38.8, **p < .000***
	50%		F (1,86) = 51.1, **p < .000***	F (1,86) = 159, **p < .000***
	75%		F (1,86) = 90.8, **p < .000***	F (1,86) = 403, **p < .000***
Sex	25%	F (1,86) = 2.74, p = .101	F (1,86) = 2.1, p = .151	F(1,86) = 2.84, p = .096
	50%		F (1,86) = 1.18, p = .28	F (1,86) = 1.77, p = .186
	75%		F (1,86) = .184, p = .669	F (1,86) = .092, p = .762
Language*sex	25%	F (1,86) = 1.81, p = .181	F (1,86) = 5.18’ **p = .025***	F(1,86)= .44, p = .507
	50%		F (1,86) = 15.9, **p < .000***	F (1,86) = .588, p = .445
	75%		F (1,86) = 29, **p < .000***	F (1,86) = .011, p = .917

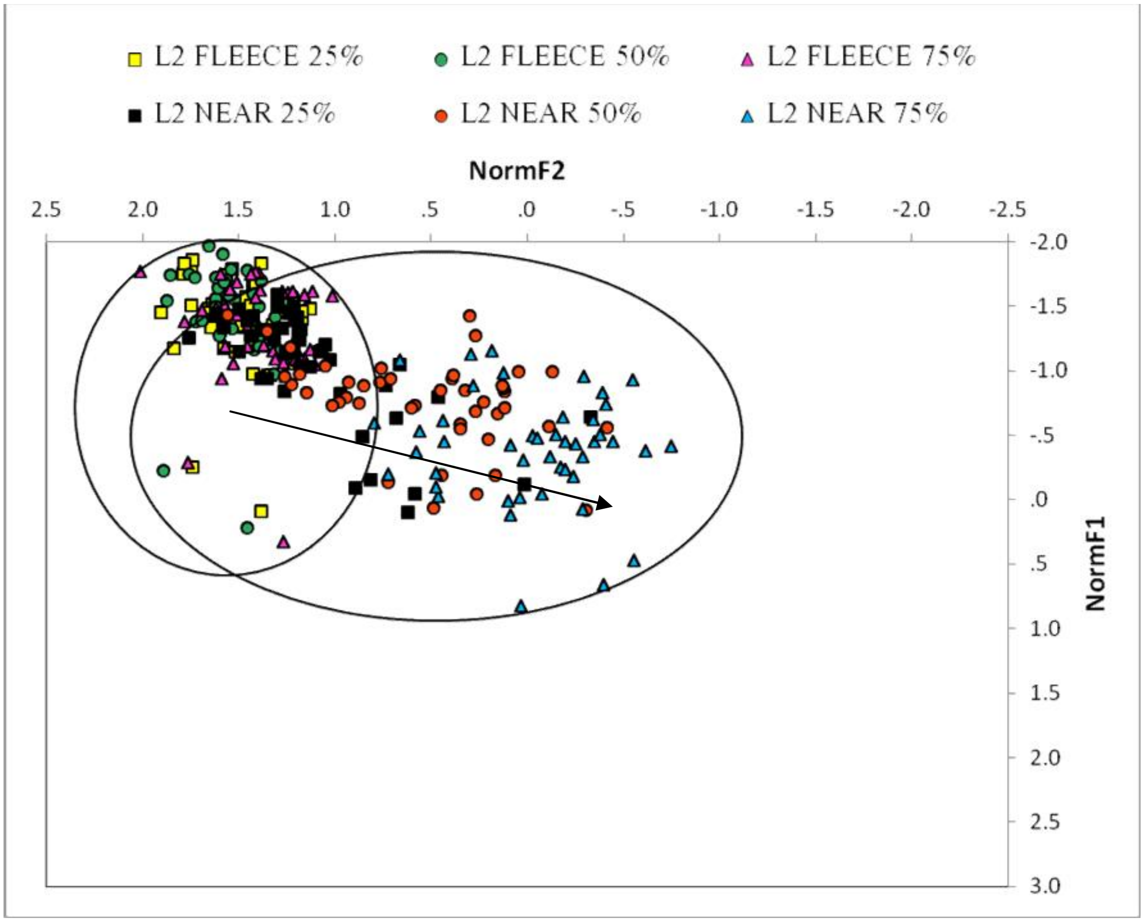

Figure 7-36: Mean normalised F1 and F2 values of SA participants' L2 productions of English (FLEECE, NEAR) contrast.

7.5 Discussion

Best's (1995) perceptual assimilation model PAM is based on the principles of the direct realist view of perceptual learning, which argues that speech perception is direct and that listeners can use the acoustic signal to infer the actual articulatory gestures (Gibson 1966, 1979). However, the constellations of articulatory gestures used for a particular sound might yield acoustic properties which differ from one language to another. Nevertheless, these sounds from the different languages may be perceived as phonologically similar suggesting that the speakers of those languages are producing phonologically similar gestures (Fowler 1996). The present chapter used the PAM framework and its perceptual assimilation types to predict the production of six English vowel contrasts. To form those predictions, the results of the PAT (chapter 5) and the identification task (chapter 6) were analysed and used to predict the productions of the members of these English contrasts by SA participants compared to those of NE and to their L1 equivalent productions. The members of these contrasts were examined to see whether they were produced with distinct or overlapping acoustic parameters.

Similarly to the identification task, the results of the production experiment showed that SA participants did not have great difficulty in producing the members of the six English vowel contrasts under investigation with distinct acoustic parameters, with the exception of English (GOAT, THOUGHT) L2 productions. Table 6-5 below presents a summary of the results of the present experiment compared to the results of the PAT and the identification task. As predicted, the table shows that all English vowels under investigation were produced with NNL acoustic properties. In general, the NE productions were longer in duration for most vowels, which might be due to the effect of the voicing of the coda in the target words /hVd/. Accordingly, though some L2 productions were qualitatively close to NE productions such as TRAP, BATH, and FLEECE, they were considered NNL. If only qualitative differences were counted, the nativity of L2 productions will be different. For example, TRAP, BATH, and FLEECE will be considered NL productions. In the following sections, the nature of these differences (quantitative or qualitative) will be discussed for the six English vowel contrasts, which will be presented and discussed, as in previous chapters, within their PAM perceptual assimilation types, i.e. TC, SC, CG, and UC.

Table 7-20: summary of the SA participants' identification and discrimination results from the previous chapter, and the production results of the English vowel contrasts under investigation compared to those of NE participants

English contrast	PAM type	Compared to NE in ID task	Identification	Discrimination	Production	Compared to NE
(TRAP, BATH)	TC	similar	(94%, 90%)	excellent	DP	(NNL, NNL)
(LOT, STRUT)	TC	different	(51%, 95%)	excellent	DP	(NNL, NNL)
(FACE, SQUARE)	SC	similar	(95%, 98%)	excellent	DP	(NNL, NNL)
(KIT, DRESS)	CG	different	(68%, 89%)	moderate	DP	(NNL, NNL)
(GOAT, THOUGHT)	CG	different	(73%, 77%)	good	OP	(NNL, NNL)
(FLEECE, NEAR)	UC	different	(94%, 73%)	very good	DP	(NNL, NNL)

7.5.1 English TC contrasts

Table 6-5 above shows that the results of the perceptual experiments of the previous chapters revealed that (TRAP, BATH) and (LOT, STRUT) English vowel contrasts were analysed as PAM TC assimilation type. The analyses of these vowels showed that the members of these contrasts were categorised into two separate SA/L1 categories (/aː/, [aːˤ]) and ([o], /a/). Additionally, these English vowels received high correct identification and discrimination, except for the moderate identification of English LOT which did not however affect its discrimination from STRUT. Based on these results and the principles of the direct realist view of speech perception, the members of these TC contrasts were predicted to be produced with distinct constellations of phonological gestures similar to those of NE. However, the phonetic detail of L2 productions were not predicted to match NE as they were predicted to be produced with distinct acoustic parameters which were NNL.

The results of the present study, summarised in Table 6-5 above, supported the predictions for the production of both English TC contrasts. SA participants managed to produce these vowels with DP using their phonological gestures which resulted in phonetic detail that shared some acoustic properties with NE and/or SA. Table 7-21 below presents a summary of the acoustic similarities and differences of the productions of SA participants' L2 compared to NE and to their L1 productions, i.e. in terms of vowel duration, F1, and F2. As can be seen, although these English vowels were phonologically assimilated in the PAT to particular SA categories, their phonetic detail does not exactly match NE or L1 productions, except the BATH vowel which matches L1 acoustically.

Table 7-21: Summary of L2 productions acoustic similarities and differences compared to NE and L1 productions (X means statistically different, √ means statistically similar)

L2 vowel	Compared to NE			Compared to L1		
	vowel duration	F1	F2	vowel duration	F1	F2
TRAP	X	√	X	√	√	X
BATH	X	√	√	√	√	√
LOT	X	√	X	X	X	X
STRUT	√	X	X	X	X	X

The English (TRAP, BATH) vowels were given high ratings of similarity to their L1 equivalents (/aː/, [aːˤ]), which was reflected in their L2 productions. As shown, quantity had a greater effect on the NNL L2 productions of English (TRAP, BATH) than quality, which suggests that there was a similarity in the phonological gestures used to produce these vowels. However, SA participants produced those vowels with a relatively similar L1 /aː/ to [aːˤ] durational ratio, which were calculated as follows: L1 (1:0.91), L2 (1:0.93), whereas the NE ratio was (1:0.6). On the other hand, English (LOT, STRUT) vowels were given low goodness of fit rating to their L1 equivalents ([o], /a/), which suggests that the members of these English vowels are produced with phonological gestures that are different from those in the L1. Unlike the previous contrast, quality had a greater role on the NNL productions of these vowels. As can be seen in Table 7-21 above the phonetic detail of L2 productions are different from both NE and L1 productions. The difficulty in the categorisation and identification of English LOT was also reflected in the difficulty in producing native-like productions. In terms of quantity, although L2 (LOT, STRUT) durations were different from their L1, the participants used a similar L1 [o] to /a/ ratio (1:1.1) for their L2 productions (1:1.18).

As discussed in the previous chapters, the TC English contrasts were mapped into SA phonetic vs. phonological categories. The phonetic categories [aːˤ] and [o] were argued to have an effect on the correct identification and discrimination of these contrasts. Similarly, the present study showed that these L1 phonetic categories had an effect on the participants' L2 productions of phonologically similar English vowels, i.e.

BATH and LOT. However, the effect of L1 phonetic as well as phonological categories does not necessarily lead to phonetically native-like productions. Instead, the L1 category effect leads to phonologically and phonetically distinct L2 productions, which would constitute the participants' L2 phonological space.

In Flege's (1995: 239) first SLM hypothesis, he argues that L2 learners perceive L2 target sounds as allophones of their L1 in a common phonological space. In line with this, the results of the present study showed that L2 productions shared some acoustic properties with English and SA equivalent productions. This finding suggests that L2 sounds, which were phonologically assimilated in the PAT to L1 categories and produced phonetically differently from both L1 and L2 sounds, are produced as allophones, i.e. perceived as similar but produced slightly differently based on the language context L1 or L2.

He also argues that for advanced L2 learners the productions of these L2 categories will eventually match their phonetic representations, by which he means they will become native-like. However, to reach this native-like target depends on the learners' motivation and the extent of native-like input. For example, the participants of the present study can be described as advanced L2 learners as many of them are English language teachers or having a degree in English. Nevertheless, their L2 input was mainly NNL from local rather than native English teachers. Therefore, their native L2 input has been limited and their motivation to be native-like is arguably limited as well. As a result, their productions of the English vowel contrasts under investigation were phonetically distinct but not native-like suggests that their main aim is to be understood. To do this, the participants used their L1 phonological and phonetic categories.

7.5.2 English SC contrast

In the PAT, Table 6-5 above shows that the English (FACE, SQUARE) diphthongal contrast was analysed as a PAM SC assimilation type since both members were categorised into a single SA vowel /eː/ with similarly high rate of similarity. I predicted in the previous chapters that the SA participants would be able to perceive the phonetic differences between these vowels, particularly the diphthongal projection. As suggested in section 6.5.2, Standard Arabic coexists with dialectal Arabic, thus the participants have access to the Standard Arabic vowel-glide sequences, which are diphthong-like sounds in Arabic. Having access to their L1 diphthongs, the participants were able to identify phonetically equivalent L2 diphthongs, i.e. (FACE, SQUARE). Therefore, they were able to highly correctly identify and discriminate the members of this contrast in the identification task. Based on the results of the perceptual experiments, L2 productions of these vowels were predicted to be NNL since they are diphthongs and phonetically they do not match directly any SA category (though they are close to the SA vowel-glide sequence /aj/). However, I predicted that the participants will attempt to produce the diphthongal projections, which they were argued to perceive, to have DP L2 productions.

The results of the present study supported the predictions for this contrast. Table 6-5 shows that L2 productions of the English (FACE, SQUARE) vowels were NNL, and they were produced with DP. Although these vowels were mapped to the SA /eː/ vowel, they were produced phonetically differently from that category. Table 7-22 below shows that L2 productions are quantitatively and qualitatively different from their L1 equivalent productions. However, L2 productions share more qualitative similarities with NE productions, but not quantitative. The reason for the quantitative difference might be related to the difference in the trajectories between NE and L2 productions, which means that NE participants hold the articulatory gestures for the on-glide and off-glide components of the diphthongs longer than SA participants do.

Table 7-22: Summary of L2 productions acoustic similarities and differences compared to NE and L1 productions (X means different, √ means similar)

L2 vowel		Compared to NE			Compared to L1		
		vowel duration	F1	F2	vowel duration	F1	F2
FACE	25%	X	√	X	X	√	X
	50%		√	√		X	X
	75%		X	X		X	X
SQUARE	25%	X	X	√	X	X	X
	50%		X	√		X	X
	75%		X	√		X	X

As discussed for the previous contrast, L2 productions of the (FACE, SQUARE) contrast were adjusted to the participants' phonological space, in which NE categories are perceived as allophones of their L1 equivalents. Therefore, these vowels were perceived as phonologically similar to the L1 monophthong /eː/, but phonetically they were producing similar on-glide and off-glide components to that of NE. However, the acoustic characteristics of these components are not the same for NE and L2 productions. For example, the components of SQUARE vowel are phonetically similar to SA [e] and [ə] phonetic categories, which were found to have a higher position in the vowel space. Thus, L2 productions of SQUARE exhibited a trajectory similar to NE but at a higher position.

As predicted, the participants were able to detect the phonetic differences between the members of this contrast as well as the differences between NE and L1 categories. Flege (1995) argues that in the case of bilinguals, being able to detect phonetic differences will help them to establish new phonetic categories for the L2 sounds which will eventually match NE categories. On the other hand, the ability to perceive these phonetic differences by the L2 learners of the present study, and producing these vowels with distinct acoustic parameters DP suggests that, similar to bilinguals, L2 and/or foreign language learners will establish new phonetic categories for these vowels. However, these categories do not match NE or L1 categories, but will eventually lead to phonetically distinct productions, which is arguably the main aim for those learners.

7.5.3 English CG contrasts

The results of the PAT in chapter 5 showed that the English (KIT, DRESS) and (GOAT, THOUGHT) contrasts were analysed as PAM CG assimilation type since the members of these contrasts were mapped into a single SA category /i/ and /oː/, respectively. English KIT and THOUGHT vowels were phonologically and phonetically mapped as closer to the L1 categories, whereas English DRESS and GOAT vowels were mapped phonologically to the L1 vowels but phonetically they were deviant. As predicted, the identification results showed that both contrasts had moderate discrimination. Additionally, the members of the latter contrast were correctly identified, with a slightly better identification of the closer to L1 vowel THOUGHT. On the other hand, the identification of the phonologically and phonetically closer to the L1 vowel KIT received poorer correct identification than its deviant vowel DRESS, contrary to what was predicted.

The better identification of the DRESS vowel and misidentification of KIT as DRESS was suggested to be the result of acoustic similarity between the DRESS vowel and the SA category /i/ rather than its equivalent category [e]. However, the results of the present study showed that there is no acoustic similarity between NE DRESS and the SA /i/ vowel. On the contrary, NE KIT productions were acoustically closer to the SA /i/ vowel. Similarly, L2 KIT productions were close to those of NE, but they were slightly shorter. As a result of this quantitative difference, L2 KIT productions were considered NNL. Table 6-5 above also shows that L2 DRESS productions, as well as, L2 (GOAT, THOUGHT) were NNL as they were quantitatively and qualitatively different from those of NE.

Based on the results of the perceptual experiments, the members of the CG contrasts were predicted to be produced with overlapping acoustic parameters OP. The present study, supported this prediction for L2 (GOAT, THOUGHT) vowels which overlapped acoustically in terms of quality and quantity. L2 productions of these vowels were monophthongal in nature, suggesting that L2 GOAT productions did not exhibit any diphthongal projection similar to that produced by NE participants. Generally, Table 7-23 below shows that L2 productions of this contrast were closer to SA monophthongal vowel productions, particularly in height. On the other hand, L2 productions of (KIT, DRESS) vowels were acoustically distinct from each other (DP). Table 7-23 below shows that L2 KIT productions were closer to NE, whereas L2 DRESS were closer to L1 [e] productions.

Table 7-23: Summary of L2 productions acoustic similarities and differences compared to NE and L1 productions (X means different, √ means similar)

L2 vowel		Compared to NE			Compared to L1		
		vowel duration	F1	F2	vowel duration	F1	F2
KIT		X	√	√	X	X	√
DRESS		√	X	X	√	√	X
GOAT	25%	X	√	X	X	√	X
GOAT	50%		X	X		√	X
GOAT	75%		X	X		√	√
THOUGHT	25%	X	X	X	√	√	X
THOUGHT	50%		X	X		√	X
THOUGHT	75%		√	X		√	√

In terms of perception, the participants perceived the English KIT and THOUGHT vowels as examples of their L1 equivalent categories (/i/, /oː/), in which, i.e. the PAT, they were predicted to perceive no phonetic differences between both L2 and L1 categories. On the other hand, the deviant DRESS and GOAT vowels were perceived as phonologically similar to the same SA categories, but phonetically the participants were able to detect some of the phonetic differences between the L2 and L1 vowels. For these contrasts, PAM predicts that a new category can be established for the deviant L2 categories, which was evident in the high correct identification of these vowels and their good discrimination from their counterpart vowels.

In terms of production, the results of the present study showed that perceiving an L2 category as phonologically and phonetically similar or nearly similar to NE does not necessarily lead to native-like productions. Fowler (1996) argues that what speakers use in production is their phonological gestures, which may yield different phonetic representations from those produced by NE. Although the participants of the present study were able to detect the phonetic differences between the members of the CG L2 contrasts, they were producing them using their L1 phonological gestures. This argument accounts for the DP production of L2 (KIT, DRESS) vowels as they are close to SA phonological vs. phonetic (/i/, [e]) categories. Also, this argument accounts for the

OP production of L2 (GOAT, THOUGHT) vowels, which were produced with the same phonological gestures used in L1 for SA /oː/ productions. In general, L2 productions of the closer-to-L1 categories were produced using the same phonological gestures used for their L1 equivalents, which yielded THOUGHT productions closer to L1 and KIT productions closer to NE (since the KIT vowel was acoustically similar to L2 and L1 productions). Since there is a great difference between the L1 vowels and the deviant English vowels, the phonological gestures used to produce the L1 vowels will work as magnets that attracts L2 productions close to it or to a half way value between L1 and NE productions. Thus, the difference between L2 productions of the deviant vowels and NE productions will be greater than the closer to the L1 vowels.

7.5.4 English UC contrast

As for the UC contrast, the English (FLEECE, NEAR) vowels were found to map to a single SA vowel category /iː/, in which FLEECE was analysed as the categorised vowel as it was mostly categorised as the L1 category with a high rating of similarity. On the other hand, the English NEAR vowel was analysed as the uncategorised member since it was highly categorised into the same SA category, and with a high rating of similarity, but not as high as the categorised vowel. These results suggested that the participants perceived the FLEECE vowel as phonologically and phonetically similar to the L1 vowel, whereas they perceived the NEAR vowel as phonologically equivalent to the same L1 vowel but phonetically deviant.

L2 productions of the uncategorised vowel were predicted to overlap with the categorised vowel since the phonetic detail of the phonological gestures used in the L1 /iː/ vowel productions do not include a diphthongal projection. However, both members were predicted to have very good identification, with a better correct identification for the categorised vowel FLEECE. As predicted, the results of the identification task showed that the members of this contrast were well identified and the categorised vowel was highly correctly identified more than its uncategorised counterpart. Contrary to what was predicted though, the participants might have been able to detect the phonetic detail of the difference between English NEAR diphthong and its SA equivalent monophthong /iː/ vowel, which is mainly in the diphthongal trajectory from a high front vowel into a mid-central vowel quality, i.e. the schwa. Thus, L2 participants produced these vowels with DP.

Additionally, the results of the present study showed that L2 FLEECE productions were NNL since the phonetic detail of English FLEECE vowel was different from the NE category. Similarly, L2 NEAR productions were also NNL. However, SA participants managed to produce the uncategorised NEAR diphthong with a diphthongal projection that was similar to NE participants'. Nevertheless, L2 NEAR diphthongal productions were at a higher position on the F1 dimension, as can be seen in Table 7-24 below which compares L2 productions to their L1 and to those of NE productions. As shown, L2 FLEECE productions qualitatively match their L1 /iː/ productions, yet they are quantitatively different as L2 productions were longer in duration.

Table 7-24: Summary of L2 productions acoustic similarities and differences compared to NE and L1 productions (X means different, √ means similar)

L2 vowel		Compared to NE			Compared to L1		
		vowel duration	F1	F2	vowel duration	F1	F2
FLEECE	25%	X	√	√	X	√	√
FLEECE	50%		X	X		√	√
FLEECE	75%		X	√		√	√
NEAR	25%	X	X	√	X	X	X
NEAR	50%		X	√		X	X
NEAR	75%		X	√		X	X

The table shows that L1 had a greater effect on the categorised vowel production, in which they produced L2 FLEECE vowel with similar phonological gestures that yielded similar phonetic detail to their L1. However, having longer L2 duration suggests that all they did was hold their articulators during FLEECE productions for a longer time than they did for their L1. On the other hand, the L1 effect on L2 NEAR productions was limited as it worked only as a magnet which attracted the phonological gestures during the production of this vowel into a higher position that is similar to the one used to produce SA /i/ and [ə], which are similar to the on-glide and off-glide components of this diphthong. Also, the articulators during the production of this diphthong were held longer in the L2 productions than in their L1 in order to produce a longer vowel that would be closer to NE productions.

Flege's (1995: 239) fifth SLM hypothesis argues that learners are unlikely to form a new phonetic category (to be native-like) for an L2 sound if it is perceived as an equivalent to a particular L1 sound, instead, L2 learners will continue to perceive that L2 sound as an example of the L1 category. Additionally, he predicts that these sounds will be similarly produced. In line with Flege's argument, the results of the PAT showed that the English FLEECE vowel was perceived as an equivalent to the SA /iː/ vowel as it received the highest number of responses into SA /iː/ and the highest rate of similarity. Also, the results of the present study showed that L2 FLEECE productions matched L1 productions in quality. On the other hand, NEAR productions were closer to NE productions than to L1, but phonetically different from both NE and L1. Flege predicts the NEAR vowel to match NE productions. However, adequate NE input and motivation to be native-like is required for the participants of this study to have native-like NEAR productions.

7.6 Summary

The present chapter examined the production of six English vowel contrasts, whose predicted phonetic similarity to their L1 equivalents was based on the PAT goodness of fit rating of the perceived similarity between both vowels, and whether an L2 contrast was mapped into one or two SA categories. In general, this study aimed to answer the following research questions:

RQ1: What are the qualitative (F1/F2) and quantitative (duration) acoustic characteristics of the participants' L2 English productions compared to NE and SA?

The first TC (TRAP, BATH) contrast was mapped into two SA categories (/aː/, [aːˁ]) with relatively high goodness of fit ratings, whereas the second TC (LOT, STRUT) contrast was mapped into two SA categories ([o], /a/) with relatively low ratings of similarity. As a result, L2 productions of the former TC contrast were phonetically closer to their L1 equivalents than the latter one. On the hand, the SC (FACE, SQUARE) contrast was mapped into a single SA category /eː/, with equally good ratings of similarity. Phonologically both vowels were similar to the L1 vowel, but phonetically both were perceived as equally deviant from that vowel. Therefore, L2 productions of the SC vowels were phonetically different from their L1 equivalent vowel.

Similarly, the UC contrast (FLEECE, NEAR) was perceived as equivalent to a single SA category /iː/, where both vowels had a high rating of similarity to that SA vowel. However, the FLEECE vowel was given a close to identical rating (6/7) suggesting that it is phonetically similar to the SA vowel /iː/. Thus, the participants were able to detect the phonetic differences between the uncategorised vowel NEAR and its phonological SA equivalent /iː/ by giving it a lower rating of similarity. As for the CG contrasts (KIT, DRESS) and (GOAT, THOUGHT), they were categorised into a single category each /i/ (including its phonetic categories [e] and [ə], which were not found to greatly differ phonetically from each others) and /oː/, respectively. Additionally, KIT and THOUGHT vowels were given higher goodness of fit ratings than their counterparts, suggesting that they are phonetically closer to the L1 categories. In production, as well as in perception, the CG contrasts were more confusing than the other contrasts, as they were phonetically produced closer to the L1 categories.

RQ2: What is the effect of the participants' L1 phonological vs. phonetic vowel space on their L2 productions?

According to these findings, an L1 effect was prevailing on L2 productions of these vowels. The same phonological gestures were used in the L1 and L2 productions resulting in phonetically close parameters. For most of these vowels, L1 phonological gestures were acting as a magnet that attracts L2 productions of the phonologically similar English vowels into a phonetically close position to that of their L1, or to a position half way between NE and L1. In terms of quantity, most L2 productions had durations shorter than NE, but longer than L1. Many reasons were suggested to explain these quantitative differences. First, NE productions tended to be longer in duration, which is perceived by SA participants as holding the phonological gestures during their L2 productions for a longer time than their L1. Thus, the participants translated the durational difference between their L1 and NE vowels to producing the English vowels longer than their L1. Another explanation was the effect of the voicing of the following consonants on the duration of the vowel, where NE productions tend to be longer before voiced consonants whereas L2 and L1 do not. Another reason related to diphthongal productions was that having L2 monophthongal productions instead, or having little diphthongal trajectory movement that does not reach its off-glide target.

Nevertheless, the effect of L1 phonetic categories was not always negative resulting in L2-NE differences. For example, the phonetic categories [aːˤ] and [o] had a great effect on the correct identification and discrimination of their English equivalents BATH and LOT vowels. Also, the SA [aːˤ] category had a great effect on producing L2 BATH vowel qualitatively similar to NE. Furthermore, having SA schwa quality, though qualitatively different from NE, had a great role on producing centring diphthongal trajectories of SQUARE and NEAR.

RQ3: After reducing the non-linguistic effects of gender and language using LOBANOV statistical normalisation procedure, which acoustic measurements do reflect phonetic differences between English, SA, and/or L2 vowels, qualitative and quantitative?

The qualitative differences between L2 productions and those of NE were attributed mostly to the effect of the participants' L1 phonological and phonetic vowel categories, which were fully analysed in chapter 4. Those included the difficulties the participants had in producing English diphthongs with a schwa as one of their components, i.e. SQUARE, GOAT, and NEAR, in which L2 productions were on a higher position that was closer to L1 schwa phonetic category. Additionally, performing a diphthongal trajectory was easier for SQUARE and NEAR diphthongs than it was for GOAT, which might be due to having an <r> as their final sound which might have prompted the participants to generalise a diphthongal production for these final <r> words. On the other hand, the target word used for GOAT did not have an <r> that can be translated into schwa, which suggests a possible effect of spelling. Furthermore, producing English diphthongs can be argued to be more demanding than monophthongs. The learners need to produce the target onset, the target offset, and produce a trajectory movement with a specific duration. In this study, the onset and offset components of these diphthongs represent higher qualities in the participants' L1 vowel space, i.e. /i/, [e], and [ə]. Therefore, qualitatively, L2 SQUARE, NEAR, and FACE diphthongs were higher than NE productions. As for duration, these diphthongs were all shorter than NE, which can be due to the shorter durations of the L1 /i/, [e], and [ə] compared to English. Also, L2 English FLEECE vowel differed qualitatively from NE productions as a result of the ‘equivalence classification (Flege 1995: 239)’ to its L1 equivalent phonological vowel /iː/.

RQ4: How do PAM perceptual predictions of the different assimilation types reflect on production?

To sum up, the PAM perceptual assimilation categories and their predicted identification and discrimination patterns from the previous chapters were a useful tool to predict L2 productions of the same contrasts. However, the perceptual patterns were not an exact match to the production patterns. The results of the analysis of the SA participants’ L2 productions of the six English vowel contrasts under investigation revealed that it was possible to infer predictions based on the perceptual patterns of these contrasts and a full analysis of the participant’s vowel system in chapter 4. Since none of the English vowels under investigation was given an identical goodness of fit rating to its SA equivalent, none of these vowels was predicted to have native-like productions. Additionally, the direct realist view of speech perception claims that speakers use their L1 phonological gestures to produce those English vowels. However, the phonetic realisations of the participants’ L1 were different from those of NE, for most vowels.

All of the above results were interpreted for the participants of the present study, who were L2 learners of English who had their English input for at least 11 years from local English language teachers. Therefore, for some of them to be understood is their ultimate aim rather than to be native–like. Also, the amount of L2 input and L2 usage for this type of L2 learners is usually limited. Therefore, the effect of L1 phonetic as well as phonological categories was prevailing throughout the three experiments. In the majority of cases, the participants’ ability to perceive the phonetic differences between English and SA equivalent vowels led to phonologically and phonetically distinct L2 productions, rather than phonetically native-like productions, which would meet the participants’ needs to be well understood. In the following chapter, the results of the three experiments of this study will be discussed thoroughly.

8 General Discussion and Conclusions

This chapter presents a general discussion of the main findings and conclusions of this thesis. It highlights the main outcomes of the perceptual and production experiments and how these are related to the existing body of literature. Also, this chapter aims to provide possible answers to the main research questions presented in chapter 1. Finally, this chapter ends with presenting the main conclusions of this thesis and some suggestions for future work.

The present thesis aimed to examine L2 perception and production of English vowels by SA Foreign Language learners, who are advanced classroom English learners with no direct native L2 input. In order to achieve this aim, theories of First Language (L1) perception and production were reviewed including the Motor Theory (MT) (Liberman, et al. 1967, Liberman, et al. 1957, Liberman and Mattingly 1985) and the Direct Realist Theory (DRT) of speech perception (Fowler 1980, 1981, 1986, 1996, Fowler and Dekle 1991), and the source-filter theory of speech production (Fant 1960), see chapter 2.

Assuming categorical perception for the L1 sounds (Casserly and Pisoni 2010, Fowler and Dekle 1991, Liberman, et al. 1957), in which L1 speakers build up perceptual boundaries between those sounds, it was argued to be more difficult to discriminate sounds within any two L1 boundaries than sounds across those boundaries. Since the target learners of the present study had a greater L1 influence than naturalistic L2 learners, because they live in the L1 country, the effect of their L1 categorical perception can be argued to overwhelmingly filter their L2 vowel perception. This argument suggests that the participants of the present study should have a greater difficulty to identify and discriminate any English/L2 categories which fall within a single SA/L1 category or boundary, than L2 categories which fall across SA boundaries. Indeed, the results of the identification study in chapter 6 showed that it was more difficult to identify and discriminate English (KIT, DRESS) and (GOAT, THOUGHT) contrasts which fall within a single SA category /i/ and /oː/, respectively, compared to English (TRAP, BATH) which fall across two SA categories /aː/ and [aːˤ], respectively. On the other hand, the categorical perception effect cannot account for the excellent identification and discrimination of the SC English contrast (FACE, SQUARE); see section 6.4.3. The members of this contrast were mapped into a single L1 category /eː/ (see section 5.4.2) but they were produced distinct from each other in the vowel space; see section 7.4.6.1.

The DRT (Fowler 1980, 1981, 1986, 1996, Fowler and Dekle 1991) assumes that what we perceive is the actual articulatory gestures which we infer from the acoustic signal. Based on this assumption, speakers of different languages use different articulatory gestures or different phasing settings (Best 1995: 193). Additionally, even when two sounds are perceived as phonologically similar across two languages, they may be produced with different acoustic outputs based on the exact constellation of gestures used and the phasings between those gestures. For example, in the PAT 80% of the English FLEECE vowel was mapped to SA /iː/, with a high goodness of fit rating 6/7, which might suggest that they are phonologically similar; and that they are produced with similar phonological gestures. However, the production of FLEECE by native English speakers was acoustically different from that of the SA productions of their L1 /iː/ vowel on two acoustic dimensions, i.e. duration and frontness, see section 7.4.6.1. The opposite case of perceiving two sounds as phonologically similar but acoustically different, is claimed by the DRT, where the same acoustic signal is produced by two different set of gestures. However, none of the English vowels examined in this study, which were given low goodness of fit, were produced with similar acoustic properties to any L1 category.

The present study aimed to account for the perception and production of English Foreign Language (EFL) learners as opposed to English Second Language learners (ESL). EFL learners have not received much attention in the literature of phonetics and phonology of L2 speech; rather this group has been the target of extensive research in L2 teaching and teacher-student contexts. Additionally, none of the present L2 perception and/or production theories account for this group of learners. Therefore, the present thesis aimed to fill this gap in the literature by conducting perceptual and production experiments on a group of EFL learners, with mid and advanced levels of English proficiency. Since, the target group of learners is not the main interest of any of the available L2 theories, some insights from the main three L2 theories: the Native Language Magnet theory (NLM), the Perceptual Assimilation Model (PAM), and the Speech Learning Model (SLM), were used to examine the perception and production of

SSBE vowels by SA EFL learners. Thus, after conducting this study, it is possible now to provide answers to the research questions which motivated this research in chapter 1: (i) To what extent current L2 models (NLM, PAM, and SLM) do account for the perceptual and production patterns of SA EFL learners? (ii) And which of these models can be extended to account for the target learners of the present thesis? And what are the required modifications, if any, to be extended?

8.1 Native Language Magnet Theory (NLM)

To begin with, the NLM theory (Kuhl 1994) is an L1 perceptual theory which was not designed to account for the L2 speech, see section 2.1. However, it was included in the analysis of this study because it has implications for L2 speech perception based on the effect of the L1, which was predicted to be prevailing for the target learners of the present study. The basic assumption of this theory is that L1 prototypes work as perceptual magnets for the members of their category and, most likely, for the members of any similar L2 categories (Kuhl 2000). Accordingly, L1 prototypes are claimed to reduce L2 adult learners' sensitivity to any L2 category that is phonetically close to an L1 prototype (Iverson and Kuhl 1995). The NLM theory predicts that L2 learners will have great difficulty perceiving L2 contrasts which do not map to different phonological categories in the L1. The results of the PAT (chapter 5) and the identification task (chapter 6) provided evidence to support this prediction.

For example, the English (KIT, DRESS) contrast is mapped mainly to a single SA phonological category /i/, and thus, the participants had difficulty identifying the KIT vowel and discriminating it from the DRESS vowel. Similarly, the participants had no difficulty identifying and discriminating English (TRAP, BATH) and (LOT, STRUT) contrasts as they were mapped in the PAT into two SA categories (/aː/, [aːˤ]) and (/a/, [o]), respectively. However, the participants may have been able to use not only their L1 phonological prototypes but also their L1 identified phonetic categories (see chapter 4), which helped them to perceive the latter English contrasts. On the other hand, the NLM theory predicts that L2 learners will have no difficulty discriminating L2 sounds which are phonetically different from any L1 prototype (Kuhl 2000). Indeed, the participants had no difficulty identifying and discriminating English (FLEECE, NEAR) and (FACE, SQUARE) contrasts though they were all mapped to single SA categories /iː/ and /eː/,

respectively. As can be seen, these contrasts have one or two diphthongs; and none of these diphthongs is phonetically similar to any L1 prototype. Therefore, the phonetic difference between the members of these contrasts was noticeable to the SA participants; and thus, they were easy to discriminate. However, the participants had some difficulty identifying and discriminating English (GOAT, THOUGHT) vowels, which might be due to having the schwa as the onset of GOAT rather than the offset. Contrary to the previous diphthongal contrasts, GOAT has a schwa as the onset whereas NEAR and SQUARE have a schwa as their offset, which replaces the /r/ in the spelling at the end of the target words used in the production task. Therefore, spelling might have had an effect on making words ending with final /r/ easier to identify as a schwa target than onset schwa which does not replace an /r/.

Based on the findings of the perceptual experiments (chapters 5 & 6), the NLM theory succeeds in explaining the L2 perceptual patterns of the SA EFL learners, particularly because it focuses on the predominant effect of the L1 prototypes. However, this theory does not take into account the phonetic and allophonic categories of the L1 and their effect on the L2. The results of the present study showed that taking the phonetic categories into consideration can explain the excellent identification and discrimination of some L2 contrasts such as (TRAP, BATH). Additionally, the NLM theory does not provide any further predictions regarding the development of L2 perception with more L2 experience. This theory does not aim to account for the production patterns of L2 adult learners. One can argue that the acoustic properties of the L1 prototypes can work as production targets for similar L2 categories. Nevertheless, even if we assume there is an effect for the native language prototypes on the production of L2 similar categories, this does not explain the NNL productions of all of the English vowels under investigation, which were also acoustically different from their equivalent L1 vowels, see chapter 7. For example, L2 TRAP and FLEECE productions were mostly mapped to particular L1 prototypes /aː/ and /iː/, respectively, but their productions were produced NNL as well as different from their L1 equivalents at lease at one acoustic dimension.

Fowler (1996) claims that what speakers perceive and produce is the phonological articulatory gestures. Taking this into consideration rather than the phonetic representation of the L1 prototypes may explain the NNL productions of the SA participants. One can argue that adult L2 learners use their L1 phonological articulatory gestures to judge the similarity between an L1 and L2 categories. Thus, rating an L2 sound as similar to a particular L1 category does not necessarily entail that they will be acoustically similar such as English FLEECE and SA /iː/. The SA participants produced the L2 sound using their L1 phonological gestures, which yielded in most cases different acoustic parameters from the ones produced by NE, such as L2 productions of English TRAP and BATH, see section 7.4.1. Thus, the NLM theory fails to extend to encompass L2 speech production and it fails to explain the mismatch of L2 perception-production patterns. For example, the NLM cannot explain the mismatch of the excellent identification and discrimination of (TRAP, BATH) and (FACE, SQUARE) and their NNL productions by SA learners. The NLM theory assumes that perceiving two sounds as phonologically similar entails that they are acoustically similar which does not coincide with the findings of this thesis. Therefore, the NLM theory cannot be extended to account for L2 learners' perception, in general, or EFL learners' perception, in particular.

8.2 Perceptual Assimilation Model (PAM)

The second and the main L2 model used in this thesis is the PAM, which was designed to specifically account for the perceptual patterns of early L2 learners with little or no L2 experience (Best 1994, 1995, 1999). This model was mainly used in the analysis of the PAT results in chapter 5 by providing a set of testable perceptual assimilation patterns, which were used to predict the identification and the discrimination patterns of the same L2 contrasts in chapter 6. The results of the present thesis showed that the assimilation patterns of the PAM were a useful tool to start with for a cross-language speech perception task. However, the descriptions of these patterns are relative and not conclusive in terms of their predicted identification and discrimination patterns. The lack of concrete quantitative descriptions for these assimilations led to relative and subjective analyses by different researchers (see section 2.3.2), including the analysis of the results of the present thesis. Nevertheless, this thesis provided the actual and concrete percentages of the cross-language assimilations of English-SA vowel systems by SA participants, which can be a useful reference for future work. Later in this chapter (section 8.5), a quantitative scale to assign PAM assimilation types to particular L2 contrasts will be presented based on the results of this thesis.

Similarly to the NLM theory and most L2 theories, this model focuses on the great effect of the L1 phonological and/or phonetic categories on those of the L2. Based on the perceived phonetic similarity to the L1 categories, PAM predicts the discrimination of these L2 categories. In the PAM-L2 (Best and Tyler 2007) version, specific predictions were given to the learnability of the L2 categories by L2 learners rather than naive foreign language listeners. Additionally, Guion et al. (2000) found in their study that PAM predicts successfully the discrimination of English consonants by early Japanese L2 learners based on the results of a cross-language mapping task (English-Japanese). Thus, they concluded that PAM can be extended to early L2 learners with a suggested revision to the Uncategorised-Categorised assimilation type (see section 2.3.4 for details).

Generally, Best and Tyler (2007: 16-17) argue for a phonetic perception of L2 by naive listeners, who perceive the phonetic detail of the L2 speech which is relevant to possible phonological functions in their L1. Best and Tyler also argue for a phonological perception of L2 by naturalistic L2 learners, who infer most of the characteristics of the L2 via natural conversations. However, they did not provide an account for classroom L2 learners, who were assumed to learn the L2 to satisfy particular educational requirements. Although, the assumption of educational requirements applies to most, if not all, classroom L2 learners, this does not mean that they do not have systematic L2 perceptual patterns which might be similar to those of the naive listeners or L2 learners. The question to ask then is what level of perception classroom learners (called EFL learners in this thesis) rely on: phonetic (similar to naive L2 listeners), phonological (similar to naturalistic L2 learners), or both phonetic and phonological. I would argue for the latter since classroom or EFL learners do have access to the phonological system of the L2 via direct learning in the classroom. However, EFL learners' phonological access is restricted to the phonemic system more than it is to the allophonic-phonetic L2 properties as they need native L2 input to develop this knowledge. Therefore, EFL learners are predicted to be sensitive to phonetic variations in the L2 speech signal, which is a feature they share with naive listeners (Best and Tyler 2007), whether these variations have a phonological function in the L2 or not, and respond accordingly.

The results of the present study showed that the SA participants were using their phonetic as well as their phonological perception. In particular, phonetic perception was evident in the participants' perceptual patterns of the English diphthongs, which were produced by a native English speaker. Although they assimilated English FACE, SQUARE and GOAT to phonological L1 categories, which were long vowels (see chapter 5), they were able to detect the phonetic differences between these L2 vowels. The detected phonetic differences are, in this case, phonologically relevant in L2 as they distinguish L2 phonological vowel categories such as (FACE, SQUARE). Additionally, the participants showed evidence of perceived phonetic distance between L2 and L1 equivalent vowels, which was evident in the goodness of fit ratings of the PAT, in which none of the English vowels received an exact match to its L1 equivalent.

Similarly to the DRT, the PAM is based on a direct perception of the articulatory gestures of speech sounds rather than the acoustic properties of the speech signal. As mentioned earlier, the articulatory gestures used in production are treated as phonological perceptual targets (Fowler 1996), whereas the acoustic signal is treated as the phonetic output that may differ for similar L1/L2 sounds (i.e. sounds which are produced with similar phonological gestures but yield differences in the acoustic signal) (Best and Tyler 2007: 26). Best and Tyler also argue that L2 learners focus on learning higher-order phonological L2 aspects such as L2 lexicon and syntax, and thus, they pay less attention to the phonetic detail of the L2. For example, Scobbie, et al. (2008) found that anterior covert gestures have negligible effects on the acoustic output of rhoticity. So, for L2 learners, focusing on the articulatory gestures in perception explains the better perception of most of the English contrasts compared to their productions.Thus, the core of the PAM accounts for the mismatch of perception-production which was found in the present thesis, i.e. excellent identification and discrimination of L2 vowels, but NNL productions.

Although the PAM accounts for L2 speech perception only, the present thesis used the perceptual results from chapters 5 & 6 to predict the production patterns of the same English vowels by the same SA participants in chapter 7. Best and Tyler (2007: 20) do support the claim that the perceptual skills of L2 learners associate positively with the degree of accuracy in L2 productions. However, the results of the production task showed that the perceptual skills of the EFL learners correlated positively with the distinct productions of the same L2 vowels, rather than with the accuracy of those productions (by which I assume Best et al. refer to native-like productions). As would be predicted from the learners' perceptual patterns, the effect of L1 phonological system (i.e. phonological articulatory gestures in PAM terms) was found to be greater on the production of the English vowels with high goodness of fit to their L1 equivalents such as the L2 productions of English TRAP vowel (see section 7.4.1.1), than on the production of those with lower goodness of fit ratings such as English NEAR vowel (see section 7.4.6.1). A possible explanation was that the phonological gestures of the L1 worked as articulatory targets in L2 productions of similar vowels, even those with low goodness of fit, and thus they worked as magnets that attracted the productions of the poor-fit vowels close to L1 productions or half way between L1 and NE productions.

In previous research, the PAM was suggested to account for the perception of early L2 learners as well as for the naïve listeners of an L2 (Guion, et al. 2000). Accordingly, PAM-L2 version was developed (Best and Tyler 2007). The difference between PAM and PAM-L2 was argued to be the level of perception of the learners, i.e. phonetic vs. phonological, respectively. With regard to the target EFL learners of the present thesis, PAM does not cover both levels of perception the EFL learners have. Thus, PAM can be a useful model to start with to examine their perceptual patterns, but it is not sufficient to explain the perception of EFL learners. The results of the present thesis showed that the PAM assimilation types examined in the PAT in chapter 5 provided predictions for the identification and discrimination of the L2 members of those types, which were supported in chapter 6, with the exception of the SC contrast (FACE, SQUARE). The SC assimilation type was assumed to show poor discrimination by naïve listeners and early L2 learners. However, the EFL learners of the present study were able to detect the phonetic differences between these vowels, which were adequate to discriminate them, which is arguably due to using their phonological knowledge from the classroom as well as their phonetic perception of the L2 productions (similar to that of the naïve listeners) which were produced by a native English speaker. Thus, the perceptual assimilation categories of PAM cannot be used to predict the perceptual patterns of EFL learners unless they take into account the EFL learners' phonetic and phonological levels of perception, which would then explain the excellent discrimination of a SC contrast such as (FACE, SQUARE).

Additionally, this study showed that using the perceptual results of a cross-language mapping task (PAT) and the results of an identification/discrimination task may predict successfully the production patterns of L2 learners. For example, highly categorised L2 sounds such as FLEECE and THOUGHT (80% and 68%, respectively) to SA/L1 /iː/ and /oː/, with high goodness of fit ratings 6/7, would be predicted to be produced more accurately (i.e. native-like) than L2 sounds such as LOT and GOAT which were not clearly categorised to one L1 category and with low goodness of fit ratings less than 4/7. Indeed, the results in chapter 7 showed that the participants succeeded to produce the similar-to-L1 categories more accurately than the deviant ones. Also, highly identified and discriminated contrasts are predicted to be produced distinct such as (TRAP, BATH) and (FACE, SQUARE), and any confusion in the discrimination is reflected in overlapped productions such as (GOAT, THOUGHT). Since PAM is a perceptual model, it does not predict the accuracy or distinctiveness of the production patterns of L2 sounds. The following section discusses the SLM, which provided some predictions of L2 speech production.

8.3 Speech Learning Model (SLM)

The third model used in the present thesis is the SLM, which was developed by Flege and his colleagues (Flege 1995, Flege, et al. 1995, Flegea, et al. 1996), see section 2.3.3. Contrary to the previous models, this model was developed to account for the perceptual and production patterns of advanced L2 learners, who had intensive native L2 input in a natural communicative context (Flege 1995). Accordingly, the target level of L2 perception of the learners of the SLM can be argued to be dominantly phonological, in which L2 learners become attuned to the phonetic detail of L2 sounds in a native-like fashion. Thus, the phonological perception of the advanced L2 learners extends beyond the phonemic inventory of the L2 to the phonotactic and allophonic variations due to their daily exposure to native L2 input. By definition, the EFL learners of this thesis do not have a daily exposure to the native L2 input; therefore, they do not have access to the phonotactic and allophonic knowledge of the L2. Thus, they only have access to the phonemic inventory of the L2 which is directly taught to them in the classroom.

The SLM assumes that the phonetic/acoustic properties of the speech signal are the basic components of speech perception. Contrary to PAM, the learners are argued to store phonetic representations of the acoustic properties of L2 sounds rather than the constellations of the articulatory gestures used to produce those sounds. Flege (1995) argues that bilinguals (I assume he means in their early stages) use their L1 phonetic categories to identify L1 and L2 sounds. With reference to the EFL learners of this thesis, the results of the identification task in chapter 6 showed that the participants may have used their full L1 vowel inventory, including its allophonic/phonetic and phonological categories, to identify and discriminate the English vowels. This was evident in their high correct identification and discrimination of English BATH, DRESS, FACE, SQUARE, and NEAR, which were phonetically similar to or had one of their components similar to SA phonetic categories (which were examined and identified in chapter 4). Contrary to the advanced L2 learners, unless FL learners have adequate motivation to develop their L2 perceptual skills by seeking native L2 input via media or internet or even living in an L2 country, I predict that they will continue to use their L1 categories, phonetic and phonological, to identify equivalent L2 categories. This suggests that, in general, FL learners show slower L2 perceptual development than advanced learners, in terms of developing new phonetic representations for new or deviant L2 categories, or even for the categories similar to the L1 ones.

Additionally, the SLM (1995: 239) claims that bilinguals keep phonetic contrast between L1 and L2 sounds which coexist in a common phonological space, by which he suggests they are “related at the allophonic level”. The results of the present study support this argument; the productions of L2 English vowels shared some acoustic properties with their L1 categories and with the target Native English (NE) categories. For example, the L2 English FACE, SQUARE, and NEAR vowels were produced as diphthongs and close to NE productions but on a higher F1 level which fits the components of these diphthongs /e/, /i/, and /ə/, in the SA phonological space. These vowels were also produced with shorter durations than those by NE speakers, but longer than their L1 equivalent long vowels.

To form a new phonetic category, the SLM argues that the learners should perceive adequate phonetic distance between the L1 and L2 sounds. Accordingly, if an L2 sound is in equivalence classification with a particular L1 sound, it is unlikely that a new phonetic category will be formed. An example which supports this claim is the perception and production of the English FLEECE vowel which is perceived equivalent to the SA /iː/ vowel. The FLEECE vowel received the highest number of categorisations and goodness of fit ratings to its SA equivalent /iː/, and in production it was produced similarly to the SA vowel. On the other hand, the participants were able to perceive the phonetic difference/differences between L1 /iː/ and English NEAR diphthong. Thus, they produced the NEAR vowel closer to NE productions, with inevitable effect of L1 on the height of the off-glide (see section 7.4.6). Therefore and similarly to PAM, the effect of L1 categories was more prevailing on those which received high goodness of fit ratings.

The case for FL learners differs from the advanced learners; I would argue that the FL learners are unlikely to perceive adequate phonetic difference between an L1 and L2 categories to form a new phonetic category for the new L2 sound unless it is phonetically salient and/or phonetically different from any L1 category, or if it is taught directly to learners in the classroom. A phonetically salient L1-L2 difference might be a durational difference, which is easier to access to learners whose L1 is a dominantly quantitative language such as Arabic (Mitleb 1981). The present study shows that diphthongisation is another example for a phonetic salient difference. A great trajectory movement was an adequate phonetic difference to perceive between L1 /eː/ and L2 SQUARE (section 5.4.2), as well as, between both L2 (FACE, SQUARE) diphthongs (section 6.4.3). Furthermore, diphthongisation may be a feature of English which receives special focus and, thus, directly taught in the classroom.

The main advantage of this model compared to the NLM theory and the PAM is that it accounts explicitly for speech production and, thus, for the perception-production link. The SLM argues that in order to have accurate L2 production, learners are required to have accurate L2 perception (again, what Flege means by accurate is native-like) (Flege 1995: 237-238). As mentioned in the previous section 8.2, the perceptual skills of the FL learners are reflected in their distinct rather than accurate productions, which fit their assumed motivation to be understood. This thesis showed that the better the identification and discrimination of an L2 contrast are, the more distinct their productions are, such as L2 (TRAP, BATH). On the other hand, L2 contrasts with some difficulty in identifying and discriminating their members, e.g. English (KIT, DRESS) and (GOAT, THOUGHT), will have difficulty in production, which is reflected in overlapping acoustic parameters.

The SLM also claims that the productions of advanced L2 learners will match their phonetic representation, which can be argued to be native-like (Flege 1995). However, one can argue that the FL learners' phonetic representations for L2 categories may or may not be native-like depending on their similarities or discrepancies from their closest L1 equivalents, if any. This means that if the L2 category is phonetically and phonologically similar to its L1 equivalent, the learners will not have any problem perceiving and producing the L2 category in a NL manner since the L1 category is an exact match to it. However, if there is any discrepancy between the L1-L2 categories, L2 productions will not be NL at least at one acoustic dimension.

The SLM claims that accurate perception of the target L2 sounds is required to have accurate L2 production (Flege 1995: 238). Implicitly, this claim indicates that L2 speech perception precedes L2 speech production, which is similar to what is assumed for the L1 development (Best 1999). This is particularly a critical indication for this thesis as it suggests that it is possible and logical to infer predictions about the L2 production patterns (which were done in chapter 7) based on the L2 perceptual results (which were extracted in chapters 5&6). Furthermore, the SLM does not assume that the L2 production development is similar to that of L2 perception, which is suggested in the model H7. This hypothesis claims that 'eventually' the production of a particular L2 category will match its phonetic representation. According to this hypothesis, the perception-production mismatch for some of the L2 vowels is justified, i.e. excellent L2 perception will be followed eventually by distinct production.

All in all, the SLM makes a good model, albeit with slightly less clear hypotheses than those of the PAM, to examine the perception and production of advanced L2 learners. However, the core of the model is based on extensive native L2 input, according to which the perceptual patterns are predicted to be NL followed by NL L2 productions. In previous literature, an attempt was made to extend the SLM to account for the learnability of early L2 learners (Guion, et al. 2000), which are arguably a closer group to the target learners of the present thesis. The findings suggests that the SLM requires further modifications and revisions to be extended to early L2 learners since it was able to predict the learnability of some L2-L1 contrasts but not all of the contrasts under investigation. Similarly, the findings of the present thesis suggest that the SLM cannot readily be extended to FL learners because it emphasises the effect of L2 input and builds its predictions based on the length of exposure to that L2 input as well as on the age of arrival AOA and age of learning AOL in the L2 environment. On the other hand, amongst the three models under investigation, the SLM is the only model which explicitly predicts the production patterns of L2 learners. Although the SLM predictions for the L2 productions were not directly applicable to FL learners' productions, this model was a starting point to build up the predictions for the FL production patterns in chapter 7.

8.4 Perception-Production Link

The gestural-based theories including the Motor Theory (MT) (Liberman, et al. 1967, Liberman, et al. 1957, Liberman and Mattingly 1985) and the Direct Realist Theory (DRT) (Fowler 1980, 1981, 1986, 1996, Fowler and Dekle 1991) argue for an intimate perception-production link. In particular, the perception and production processes are assumed to share a similar set of invariants, which refer to the perceived actual/intended articulatory gestures which are also claimed to be used in production, which are arguably mastered at later stages, see section 2.1. Assuming this intimate relationship has a great impact on the phonetic and phonological organisation of the language (Best 1995: 172). Furthermore, it has been argued that speech perception precedes production during the development of the L1 (Best 1999). With regard to the L2, the PAM (Best and Tyler 2007: 20) suggests that the perceptual skills of L2 learners correlate positively with the accuracy in L2 productions, as discussed in section 8.38.2 above. Similarly, the SLM (Flege 1995) claims that the L2 learners, ultimately, aim to map their production to their phonetic representations of the L2 target categories. The PAM suggestion and the SLM hypothesis implicitly suggest that L2 perception precedes L2 production similarly to the L1. In light of the results of the present thesis, what can we infer about the perception-production link for the FL learners compared to the advanced L2 learners and to the L1?

This thesis showed that the same can be assumed for the FL learners of the present study, i.e. their L2 perceptual patterns (chapters 5&6) predict their productions (chapter 7). However, the results of the present study showed that the perceptual patterns of the FL learners were positively correlated with the distinctiveness of the L2 productions rather than the accuracy of those productions compared to those of the NE, which can be argued to be due to the lack of native L2 input. This finding suggests that indeed, similarly to L1 and L2, the FL perception precedes production. The present study showed that the perception (identification and discrimination) of the English vowels under investigation were much better than their productions. Furthermore, the perceptual results for some vowels such as (FACE, SQUARE) were very close to native-like results. Thus, the advanced perception over production is evidence that the FL Perceptual development precedes the production development.

To support this argument for the FL, I will provide a hypothetical classroom context to teach the English /p/ sound to FL Arab learners, which is derived from my own teaching experience. Figure 8-1 below presents a schematic representation of the perception-production link, which I suggest for the FL learners. Usually a classroom context starts with an input, which is provided mostly by non-native teachers. The teacher will introduce the English /p/ sound by writing it, pronouncing it, and maybe giving some real lexical items with a /p/ sound. Then the learners process this input in terms of their L1 categories (whether articulatory (PAM) or phonetic (SLM)) till they find a close match, e.g. /b/. Finally, they try to produce the L2 category similarly to its closest L1 equivalent, i.e. /p/ → /b/. However, if any further instructions provided by the teacher on the pronunciation of the English /p/, which involve unfamiliar phonetic cues (aspiration in this case [p^h]); they are predicted to be processed by the learners. If the new phonetic cues are salient and detectable, the learners may attempt to produce the L2 sound according to the instructions, if not, most of the learners will use the L1 variant as long as it does the job.

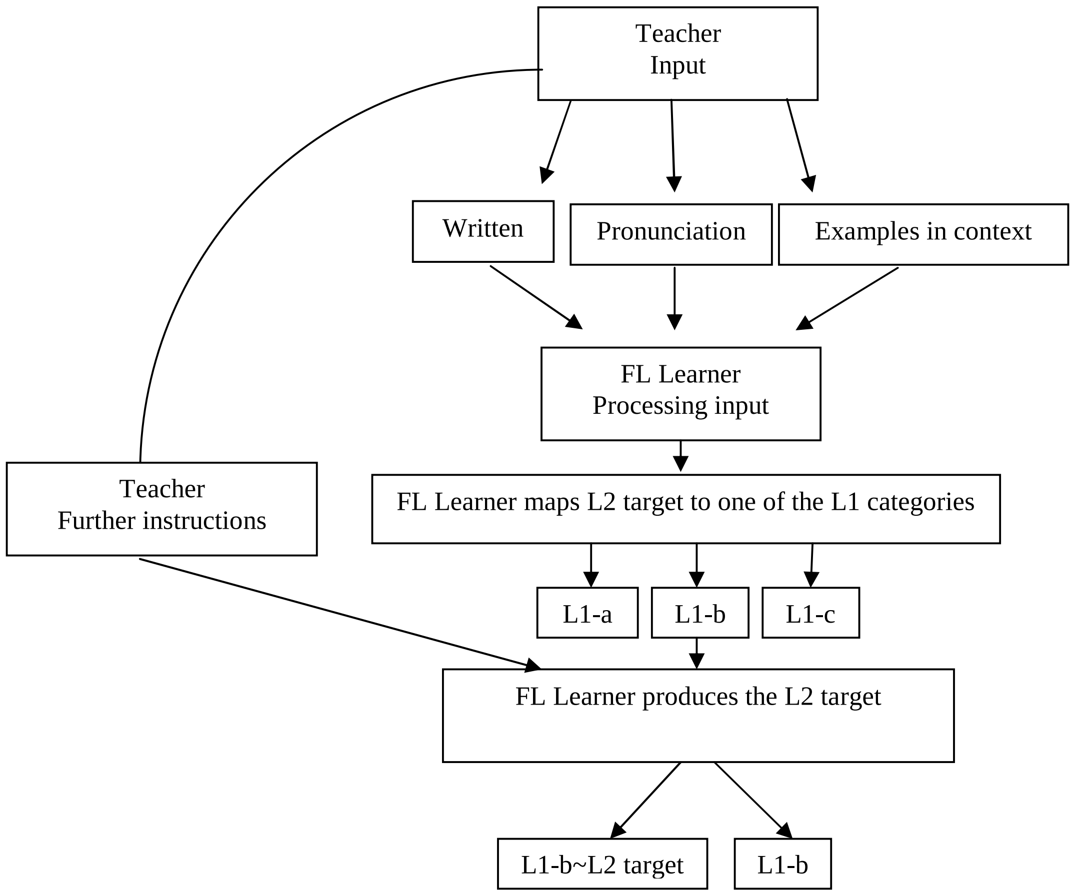

Figure 8-1: Schematic representation of the perception-production link of a hypothetical FL classroom situation

As can be seen from Figure 8-1 above, the teacher is the source for the phonological input of the L2, which is given to learners directly. Any further instructions provided by the teacher may be the source for the phonetic detail of the L2 categories. In theory, the phonetic input should be accurate and based on native data. However, what happens in practice is that the non-native teacher has the knowledge but lacks the native accuracy. Accordingly, erroneous phonetic representation may result from the lack of accuracy. At this point, further investigations within an educational framework are required to explore possible ways of enhancing the phonetic accuracy of the instructions provided by FL teachers.

In sum, the perception-production link is similar to advanced L2 learners and to the L1 in two ways. First, FL perception precedes production. The FL learners perceive any L2 input introduced by the teacher and process it, in order to produce it with varying effects of their L1. The extent of the L1 effect relies on the accuracy of the phonetic instruction provided by the teacher and the learners' motivation to develop their L2 by seeking native L2 input. Second, the perceptual skills of the FL learners are realised in the production of distinct L2 categories, i.e. being understood even with a degree of foreign accent.

8.5 Foreign Language Model (FLM)

As mentioned in section 8.2 above, the perceptual assimilation types of the PAM were a useful tool to start with to predict the identification and discrimination of L2 contrasts. However, the literature using PAM is not clear enough to show how and on what basis PAM assimilation types are assigned to particular L2 sounds, and this made the application of those types subjective. A contribution of this thesis would be providing a quantitative scale to the PAM assimilation types to be able to decide which L2 sounds map to which PAM assimilation type. The scale is suggested based on the results and analyses of the perceptual experiments in chapters 5 & 6. This thesis makes an example of a long journey trying to assign those types and explain how and why they were assigned to the L2 contrasts under investigation. Therefore, I am in a position which enables me to present a scale which can be used in future research to help researchers make decisions about PAM perceptual assimilation types. The scale assumes a cross-language mapping task with goodness of fit ratings and an identification task. The scale presents the percentages of categorisations of L2 sounds to their L1 equivalents, the percentages of goodness of fit ratings, and the percentages of correct identification of the members of the L2 contrast. Based on the association of the three percentages, a decision can be made about which PAM type to assign to a particular L2 contrast. Since, the scale is based on the results of the present study, only the examined PAM contrasts will be presented: TC, SC, CG, and UC. The suggested scale is based on percentages calculated for categorisation, goodness of fit, and identification from the three main experiments of this study in chapters 5-7.

- First, for an L2 contrast to be assigned a TC assimilation type:
  - Categorisation: over 50% of both members of the L2 contrast should be categorised to two separate L1 categories
  - Goodness of fit: both members should be given over 70% goodness of fit ratings (i.e. 5/7 in this study)
  - Identification: both members should have over 90% correct identification
- For an L2 contrast to be assigned a SC assimilation type:
  - Categorisation: over 50% of both members categorised to a single L1 category
  - Goodness of fit: both members over should be given 70% goodness of fit ratings (i.e. 5/7 in this study), and mostly they may have similar ratings.
  - Identification: depends on the perceived phonetic difference between the SC members, the bigger the difference the higher the percentage of correct identification.
- For an L2 contrast to be assigned a CG assimilation type:
  - Categorisation: over 30% of one member and less than 30% of the other member categorised to a single L1 category
  - Goodness of fit: the member with more categorisations should have over 64% (4.5/7) ratings which should be higher than the ratings for the other member.
  - Identification: the member with more categorisations should have over 75% correct identification; the other member should be less but not less than 50%.
- For an L2 contrast to be assigned a UC assimilation type:
  - Categorisation: over 75% of the categorised L2 sound should be categorised to an L1 category, over 50% of the uncategorised L2 sound should be categorised to the same L1 category.
  - Goodness of fit: the categorised member should have high ratings over 85% (6/7); the uncategorised one should have high, but lower than the categorised, rating 70% (5/7).

- o Identification: the categorised member should have high correct identification over 90%; the uncategorised one should also have high correct identification, but lower than the categorised, over 70%.

As can be seen, one can predict the identification of L2 contrasts based on the categorisation and goodness of fit ratings of the members of those contrasts. For example, it was possible to predict very good identification (73%) and excellent identification (94%) for the members of the UC contrast (NEAR, FLEECE), respectively (see section 6.4.6), based on their categorisations and goodness of fit ratings: NEAR (62% with 5/7 rating) and FLEECE (80% with 6/7 rating) (see section 5.4.6). Another contribution of this thesis would be an extension to PAM and this scale; one can also predicts the production of the members of the L2 contrasts DP or OP. For example, this thesis showed that a confusion or misidentification of one member as the other of more than 10% would result in overlapping productions. This was particularly evident in the production of (GOAT, THOUGHT) CG contrast which were misidentified 23% and 14%, respectively.

The main aim of the present thesis was to fill a gap in the literature by examining the perceptual and production patterns of FL learners, with a minimum of ten years of learning, who had received less attention within L2 phonetics and phonology research than early and advanced L2 learners. The findings and the analysis of the results of this thesis provided an informative background to develop a Foreign Language Model (FLM), which can account for the perceptual and production patterns of FL learners. This model aims to complement the important contributions of Best's PAM and Flege's SLM to understanding L2 speech perception and production. It is an attempt to bring new insights into the perception and production of L2 speech. Table 8-1 below summarises the main principles of the suggested FLM based on the analysis of the results and findings of the present thesis. These principles represent the first attempt to account for the perception and production of FL learners. So, the FLM principles are subject to further modifications and/or expansion based on the results of future research on the same group of learners from different L1 backgrounds, or comparing the FL learners to early and advanced L2 learners.

Table 8-1: The principles of a suggested FLM based on the results of the present thesis.

	FLM principles
P1	The target learners of the FLM are the FL learners who had their L2 experience in a classroom context for many years (e.g. ten years or more)
P2	The FL learners have mainly phonological perception (mainly phonemic unless the learners receive further phonotactic/allophonic instructions). They also have phonetic perception similar to that of the naïve listener, i.e. they can detect any phonetic differences between L1 and L2 sounds or between two L2 sounds if the differences are noticeable, or learners can be trained to respond to these differences via direct instruction.
P3	The FLM accounts for speech perception as well as speech production. FL speech perception precedes FL speech production, and the perceptual skills are reflected in the distinctiveness of L2 productions (DP, or OP if the perceptual skills were poor).
P4	A new phonetic category can be established for an L2 sound if it differs phonetically from the closest L1 sound, and if the FL learners could be at least taught some of the phonetic differences between the L1 and L2 sounds.

In this section, the principles of the FLM are discussed and explained in detail in an attempt to highlight the main L2 theoretical issues. The first principle (P1) specifies the target learners of the model, who are experienced classroom FL learners. This principle specifies the first and most important distinction between the FLM and the PAM, which is interested in naive or early L2 learners, and the SLM, which is interested in naturalistic advanced L2 learners. To summarise some of the well known facts about FL learners: (i) the FL learners are classroom learners who learned the L2 to satisfy specific educational requirements (Best and Tyler 2007: 16), (ii) the FL group encompasses a greater number of learners than early and advanced L2 learners, and thus, this leads to greater variability within and across learners from different L1 backgrounds, (iii) this group of learners does not have a daily exposure to native L2 input. According to these facts, the FL learners are closer match to the descriptions for early L2 learners or even naïve L2 listeners, which are accounted for in PAM-L2 and PAM, respectively. The only exception would be the non-native classroom L2 experience of the FL learners, which leads to an important distinction between the target learners of the PAM and FLM, i.e. the level of perception in the L2, which is discussed in section 8.2 above.

The level of L2 perception leads the discussion to the FLM second principle (P2), in which I argue that the FL learners share with the advanced learners of SLM their phonological perception and share with the naive listeners of PAM their phonetic perception. Contrary to the advanced learners of SLM, the FL learners can be argued to have access to the phonological knowledge via direct instruction in the class room and not via natural conversation. Thus, the FL learners' phonological knowledge can be mainly phonemic, whereas the advanced learners have access to native L2 input via which they infer the phonotactic and allophonic knowledge. However, such detailed phonological knowledge may be provided by the teacher in the classroom, but it is unlikely to be native-like. Similarly to the naive listeners, due to the lack of native L2 input, the FL learners can be sensitive to the phonetic variability in the speech signal whether this variability has a phonological function or not. Therefore, FL learners are predicted to detect phonetic differences between L1 and L2 sounds or between two L2 sounds if these differences are salient and detectable, or if they are provided as direct instructions in the classroom (see Figure 8-1 above).

The FLM third principle (P3) explicitly links FL learners' speech production to their perception. It assumes that the development of FL speech perception precedes speech production, which coincides with the description for the L1 (Best 1999) and the SLM claim, which argues that accurate perception precedes accurate production (Flege 1995: 238). However, the perceptual skills of FL learners lead to distinct productions (DP) rather than accurate/native-like productions, which was evident in the production results in chapter 7. The L2 contrasts which showed excellent identification and discrimination patterns such as (TRAP, BATH) and (FACE, SQUARE), were produced with Distinct acoustic Parameter (DP), whereas the L2 contrasts which showed some identification and discrimination difficulties such as (GOAT, THOUGHT), were produced with Overlapping acoustic Parameters (OP). This distinction between FLM and SLM originates from the difference in the ultimate goal of their target learners. The SLM is interested in the "ultimate attainment of L2 pronunciation" (Flege 1995: 239), which is accurate pronunciation, whereas the main goal for the FLM target learners is to be understood, i.e. to have distinct productions.

The fourth principle (P4) suggests that category formation is possible for FL learners. In the early stages of L2 learning, I predict a great variability in the productions of L2 taught categories. After a few years of practice and classroom instruction, a phonetic representation might be established for that category within the L1 phonological space. However, if the L2 category was taught as an exact equivalence to a particular L1 category (which is similar to the SLM H5 ‘equivalence classification’ (Flege 1995: 239)), the FL learners are not predicted to develop a separate phonetic representation for that L2 category; and it will be perceived and produced equivalent to the L1 category. On the other hand, if the learners were taught that a particular L2 category is different from any L1 category, and they were given instructions how to produce it by highlighting some of its phonetic properties, it will be eventually produced phonetically different from any L1 category. However, the FL productions of the different L2 category will not necessarily match native L2 productions as the instructions may not be native-like. In sum, the principles of the FLM offer a great range of empirical and theoretical implications for further research, investigation, and comparison, which are going to be the subject of my future research for the next few years.

8.6 Summary and Future Work

The present thesis examined the perception and production of SSBE vowels by SA FL learners. The FL learners have received little attention in the literature of L2 phonetics and phonology due to the lack of native L2 input. In order to account for this group, theories of L1 perception (MT and DRT) and L1 production (the source-filter theory) were reviewed in chapter 2. Additionally, the current L2 learning models including the NLM theory, the PAM, and the SLM were reviewed and compared in order to identify the status of the FL learners compared to the target learners of these models. Finally, the relevant literature describing the qualitative and quantitative characteristics of the L1/SA and the L2/SSBE of the target learners of this study were reviewed, followed by a review of the L2 studies involving Arabic as the L1 and English as the L2 to identify the areas of difficulties Arab learners of English have, generally, in their perceptual and production patterns of English.

Chapter 2 showed that indeed the FL group of learners is different from the target learners of any of the current L2 learning models. Thus, this research made a critical contribution to the field of L2 perception and production studies. In particular, this study examined the perception and production patterns of the SA learners, whose L1 vowel system differed from that of SSBE. The vowel system of the dialect of SA was not fully examined in any previous study. Therefore, the present study contributed to the Arabic literature by conducting an acoustic analysis on the full vowel system of SA in chapter 4. The analysis of this chapter had a great effect on the analysis of the results of the rest of the thesis in two ways: first, it provided a quantitative description of the vowel system of the L1. Second, the analysis of chapter 4 determined the phonological vs. phonetic vowel categories of SA, which had an effect on the interpretation of the L2 results in the following chapters.

The perceptual Assimilation Task (PAT) in chapter 5 and the identification task in chapter 6 were analysed based on the PAM assimilation types and their predicted identification and discrimination patterns. These chapters showed that the assimilation patterns from the PAT managed to predict the identification and discrimination patterns of most of the PAM categories, except for the SC (FACE, SQUARE) contrast which was predicted by PAM to show poor discrimination. However, the results in chapter 6 showed that this contrast had excellent identification and discrimination, see section 8.2. The perceptual results of chapters 5 and 6 were successful in predicting the production patterns of the FL learners in chapter 7, which suggests a possible extension of PAM to account for L2 speech production in addition to L2 speech perception. It is important to highlight that the design of the PAT was demanding by asking the participants to think in their L1 and to force them to assimilate and judge English vowels as equivalent to their L1 at the same time, while having 14 L1 options. In particular, the participants struggled when asked to assimilate dissimilar sounds which they believed do not fit any L1 category such as diphthongs. Therefore, further work is needed to enhance this task and check its reliability.

The present thesis also contributed to the literature on the perception-production link in two ways: first, this study showed that similar to the L1 and L2, the FL perception precedes production. Second, this study argues that the perceptual skills of the FL learners are reflected in their productions. Excellent perceptual skills were found to manifest as DP productions, whereas the poor perceptual skills were found to manifest as OP productions, see section 8.4.

Furthermore, the description of the FL learners of the present study is a better match to the PAM learners since naïve listeners or early L2 learners have a great effect of the L1 on their L2 perception, which is similar to the FL learners. Furthermore, similarly to naïve listeners, the FL learners have no native L2 input. On the other hand, similarly to the SLM advanced learners, the FL learners are taught directly in the classroom some of the phonological and sometimes the phonetic detail of the L2 by their local teachers. However, the perceptual and production hypotheses of the SLM cannot be directly applied to the FL learners due to the great difference between the FL and advanced L2 learners in terms of the level of their L2 perception and their ultimate production aim; see section 8.3. Therefore, the FL learners do not match the target L2 learners of the PAM or the SLM. Thus, the current L2 models cannot account for the perceptual and production patterns of the FL learners' speech.

The analysis of the results and the findings of the present thesis set a background to suggest a FLM (in section 8.5) which accounts for the perception and production patterns of FL classroom learners. This study provided the main principles of the FLM, which covers issues like: the target learners of the FLM and how they differ from the target learners of the current L2 models, the linguistic perception level, the perception-production link of the FL learners, and category formation. This model will be the subject for further research and examination in the future for example to cover FL learners from different L1 backgrounds and the development of FL speech perception and production. Additionally, the results of this study and similar studies on FL learners will be compared with each other and with equivalent studies on early and advanced L2 learners.

This thesis contributes to FL perception and production research by developing the FLM which can account for this group of learners; in particular it takes into account the peculiarity of this group and how it differs from early and advanced L2 learners. Additionally, this study identified the areas of difficulties the SA EFL learners may encounter in perception and production, which has a great impact on EFL teaching by encouraging FL educational research to test possible ways to enhance the perception and production of FL learners within the classroom context.

Finally, the future research based on this thesis can be in two directions. The first direction is Arabic L1studies. In particular, the vowel system of other Arabic dialects needs to be examined and compared. The majority of Arabic studies focuses on the high and low vowels and ignores dialectal Arabic mid vowels. This thesis showed that the status of mid vowels in SA had a great effect on the participants' L2. Thus, a phonological analysis of Arabic mid vowels including the distribution of mid vowels as opposed to vowel-glide sequences is needed. This thesis showed that there are quantitative and qualitative distinctions between Arabic short-long vowels, which is predicted to have had an effect on the participants L2. Thus, the role of quality and quantity in the V-VV distinctions needs to be examined. This thesis showed some interesting findings which need further examination in the future. For example, the results of CH7 showed that English vowels were in general longer in duration than SA ones, whereas the opposite was expected for SA long vowels, i.e. SA long vowels were expected to be longer than NE ones. Also, L2 productions of the SA participants were not as long as expected before voiced coda (/d/ in /hVd/), which can be due to having no effect of coda voicing on vowel duration in their L1 and thus in their L2. The present study did not include voiceless codas to be compared to L2 productions before voiced codas; however, if coda voicing has an effect on the participants' L2 productions, I predict L2 vowels to be shorter before voiceless codas, which needs further research to explore and to compare to NE productions. Another unexpected finding was that SA long /oː/ and /iː/ vowel productions exhibited greater diphthongization than English long vowels THOUGHT and FLEECE. Diphthongisation was not examined for Arabic vowels in any previous research, and the findings of this study shows interesting patterns worth of further analysis in the future.

The second direction of future research is L2 studies, in which the principles of the FLM will be the subject of further examination and analysis. For example, the phonological and phonetic levels of FL perception will be monitored and tested experimentally for a group of classroom learners, in order to determine the effect of classroom detailed phonetic instructions on category formation for sound categories such as diphthongs. After that, the productions of the FL learners will be examined to determine the development of FL speech production compared to FL perception. Additionally, the results of the present thesis and any future work applying the FLM will be compared to early and advanced L2 learners who had spent some time in an English speaking country. This comparison will highlight and confirm the differences between the three groups of learners in terms of the level of their L2 perception, i.e. phonological, phonetic, or both. These studies will also examine the effect of having native L2 input vs. the lack of native L2 input on the perception and production of learners.

Furthermore, other difficulties Arab learners of English may have will be examined, including segmentals and suprasegmentals. Finally, the following step will be related to the present thesis, the FL productions are planned to be presented to NE listeners in an intelligibility task to have a quantitative measure of how close those productions are to native-like productions, particularly diphthongs, and to examine the effect of the classroom phonetic instruction on the FL productions.

9 References

Abou Haidar. 1994. Norme linguistique et variabilité dialectale: analyse formantique du système vocalique de la langue arabe. *Revue de Phonétique Appliquée.* 110: 1-15.

Adank, P., R. Smits and R. V. Hout. 2004. A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America.* 116: 3099-3107.

Al-Ani, S. and M. El-Dalee. 1984. Tafkhim in Arabic: the Acoustic and Psychological Parameters. In proceedings of the *Tenth International Congress of Phonetic Sciences,* 385-389. Dordrecht.

Al-Tamimi, J. 2007 a. Static and Dynamic cues in Vowel Production: a cross dialectal study in Jordanian and Moroccan Arabic. In proceedings of the *The 16th International Congress of Phonetic Sciences (ICPHS),* 541-544. Saarbrücken, Germany.

Al Masri, M. and A. Jongman. 2003. Acoustic Correlates of Emphasis in JordanianArabic: Preliminary Results In proceedings of the *Texas Linguistics Society* 96-106. Somerville.

Albashir, A. 2008. Production and Perception of Libyan Arabic vowels. Ph.D. dissertation, University of Newcastle Upon Tyne.

Allatif, O. 2008. Contrôle des corrélats temporels et spectraux de la quantité vocalique: de l'arabe syrien de l'Euphrate au français de Savoie. Ph.D. dissertation, Préparée à l'ex-Institut de la Communication Parlée (GIPSA-LAB) dans le cadre de l'Ecole Doctorale: « Langues, littératures et sciences humaines ».

Allatif, O. and C. Abry. 2004. Adaptabilité des paramètres temporels et spectraux dans l'opposition de quantité vocalique de l'arabe de Mayadin (Syrie). *XXVèmes Journées d'Etudes sur la Parole.* 13-16.

Almbark, R. A. 2008. A Sociophonetic Study of Emphasis in Syrian Arabic. Ph.D. dissertation, The University of York.

Almbark, R. A. 2011. Production and Perception of SSBE Vowels by Syrian Arabic Speakers. In *Achievements and Perspectives in SLA of Speech: New Sounds 2010,* ed. Magdalena Wrembel, Małgorzata Kul and K. Dziubalska-Kołaczyk, 15-26. Peter Lang: International Verlag der Wissenschaften.

References

Andruski, J. and T. Nearey. 1992. On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables. *Journa of the Acoustical Society of America.* 91: 390-410.

Aoyama, K., J. E. Flege, S. G. Guion, R. Akahane-Yamada and T. Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics.* 32: 233–250.

Arthur, A. and S. House. 1961. On Vowel Duration in English. *The Journal of the Acoustical Society of America.* 33: 1174-1178.

Asfoor, M. A. 1982. Difficulties English Speakers Encounter in Arabic Phonology. Ph.D. dissertation, The University of San Francisco.

Best, C. 1994. Development of Language-specific Influences on Speech Perception and Production in Pre-verbal Infancy. .

Best, C. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research* ed. Strange, W., 171–204. Baltimore: York Press.

Best, C. 1999. Development of Language-Specific Influences on Speech Perception and Production in Pre-verbal Infancy. In proceedings of the *International Congress of Phonetic Sciences*, 14, 1261–1263. Department of Linguistics, University of California at Berkeley.

Best, C. T., G. McRoberts and E. Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America.* 109: 775-794.

Best, C. T. and M. D. Tyler. 2007. Nonnative and Second Language speech perception: Commonalities and Complementarities. In *Second Language Speech Learning: The role of language experience in speech perception and production* ed. Munro, M. J. and O. S. Bohn, 13-34. Amsterdam: John Benjamins.

Boersma, P. and D. Weenink. 2009. Praat: doing phonetics by computer (Version 5.1.14) [Computer program]. http://www.praat.org/. 27-08-2009.

Bohn, O.-S. and J. E. Flege. 1992. The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition.* 14: 131-158.

Bohn, O. S. and J. E. Flege. 1990. Perception and Production of a New Vowel Category by Adult Second Language Learners. In proceedings of the *New Sounds 90: Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech* 37-56. Amsterdam: University of Amsterdam .

Browman, C. P. and L. Goldstein. 1989. Articulatory Gestures as Phonological Units. 6: 201-251 .

Browman, C. P. and L. Goldstein. 1992. Articulatory phonology: an overview. *Phonetica.* 49: 155–180.

Card, E. 1983. A Phonetic and a Phonological Study of Arabic Emphasis. Ph.D. dissertation, Cornell University.

Casserly, E. D. and D. B. Pisoni. 2010. Speech Perception and Production. *Wiley Interdisciplinary Reviews: Cognitive Science.* 1: 629–647.

Cebrian, J. 2009. Exploring the roles of instruction and word familiarity in L2 vowel identification. In proceedings of the *Phonetics Teaching & Learning Conference PTLC2009,* UCL.

Chladkova, K. and S. Hamann. 2011. High Vowels in Southern British English: /u/-fronting does not result in merger. *ICPhS.* 17: .

Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. London: Harper & Row.

Cowell, M. 1964. *A Reference Grammar of Syrian Arabic*. Georgetown University Press.

Cruttenden, A. 2008. *Gimson's Pronunciation of English*. Hodder Education.

Davidson, l. 2011. Phonetic and phonological factors in the second language production of phonemes and phonotactics. *Language and Linguistics Compass.* 5: 126–139.

Davis, S. 1995. Emphasis Spread in Arabic and Grounded Phonology. *Linguistic Inquiry*. 26: 465-498.

Denes, P. 1955. Effect of Duration on the Perception of Voicing. *Journal of the International Phonetic Association.* 27: 761-764.

Denes, P. and E. Pinson. 1993. *The Speech Chain: The Physics and Biology of Spoken Language*. New York: W.H. Freeman and Company.

Derwing, T., M. Rossiter, M. Munro and R. Thomson. 2004. Second Language Fluency: Judgments on Different Tasks. *Language Learning.* 54: 655–679.

Deterding, D. 1997. The Formants of Monophthong Vowels in Standard Southern British English Pronunciation. *Journal of the International Phonetic Association.* 27: 47-55.

Diehl, R., B. Lindblom, K. Hoemeke and R. Fahey. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics.* 24: 187–208.

Diehl, R. L., A. J. Lotto and L. L. Holt. 2004. SPEECH PERCEPTION. *Annu. Rev. Psychol.* 55: 149–79.

Eimas, P., J. Miller and P. Jusczyk. 1987. On infant speech perception and the acquisition of language. In *Categorical Perception: The Groundwork of Cognition,* ed. Harnad, S., 161-195. Cambridge University Press.

Escuderoa, P. and K. Chládková. 2010. Spanish listeners' perception of American and Southern British English vowels. *Journal of the Acoustical Society of America.* 128: .

Fabricius, A., D. Watt and D. Johnson. 2009. A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change.* 21: 413–435.

Fabricius, A. H. 2007b. Variation and change in the TRAP and STRUT vowels of RP: a real time comparison of five acoustic data sets. *Journal of the International Phonetic Association.* 37: 293-320.

Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton & Co.

Ferragne, E. and F. Pellegrino. 2010. Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association.* 40: .

Flege, J. 1979. Phonetic interference in second language acquisition. Ph.D. dissertation, Indiana University.

Flege, J. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics.* 15: 47-65.

Flege, J. 1995. Second Language Speech Learning: theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research,* ed. Strange, W., Timonium, MD: York Press.

Flege, J. and I. MacKay. 2010. "Age" effects on second language acquisition. In proceedings of the *6th International Symposia on the Acquisition of Second-langauge Speech,* Poznan, Poland.

Flege, J., I. Mackay and D. Meador. 1999. Native Italian speakers' perception and production
of English vowels. *Acoustical Society of America.* 106: 2973-2987.

Flege, J. and R. Port. 1981. Cross-language phonetic interference: Arabic to English. *Language and Speech.* 24: 125-146.

Flege, J., N. Takagi and V. Mann. 1995. Japanese Adults can Learn to Produce English /r/ and /l/ Accurately. *Language and Speech.* 38: 25-55.

Flege, J. E., E. M. Frieda, A. C. Walley and L. A. Randazza. 1998. Lexical Factors and Segmental Accuracy in Second Language Speech Production. *Studies in Second Language Acquisition.* 20: 155-187.

Flegea, J., N. Takagi and V. Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Acoustic Society of America.* 99: 1161-1173.

Flynn, N. 2011. Comparing Vowel Formant Normalisation Procedures. *York Papers in Linguistics.* Series 2: .

Fowler, C. 1980. Coarticulation and theories of extrinsic timing. *Journal of Phonetics.* 8: 113-33.

Fowler, C. 1981. Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research.* 46: 127-39.

Fowler, C. 1986. An event approach to the study of speech perception from a directrealist perspective. *Journal of Phonetics.* 14: 3–28.

Fowler, C. 1996. Listeners do hear sounds, not tongues. *Acoustic Society of America.* 99: 1730-1741.

Fowler, C. 2008. The FLMP STMPed. *Psychonomic Bulletin & Review.* 15: 458-462.

Fowler, C. and D. Dekle. 1991. Listening with eye and hand: cross-modal contributions to speech perception. *Journa of Experimental Psychology: Human Perception and Performance.* 17: 816-828.

Gairdner, W. H. T. 1925. *The phonetics of Arabic*. London: Oxford University Press.

Gay, T. 1968. Effect of Speaking Rate on Diphthong Formant Movement. *Journal of the Acoustical Society of America.* 44: 1570-1573.

Gay, T. 1970. A perceptual Study of American English Diphthongs. Ph.D. dissertation, City University of New York.

Gibson, J. J. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilichinskaya, Y. D. and W. Strange. 2010. Perceptual assimilation of American English vowels by inexperienced Russian listeners. *The Journal of the Acoustical Society of America.* 128 80-85.

GraphpadSoftware. 2012. Analyze a 2x2 contingency table. http://www.graphpad.com/quickcalcs/contingency1.cfm. April.

Guion, S., J. Flege, R. Akahane-Yamada and J. Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journa of the Acoustical Society of America.* 107: 2711-2724.

Halberstama, B. and L. J. Raphael. 2004. Vowel normalization: the role of fundamental frequency and upper formants. *Journal of Phonetics.* 32: 423-434.

Hawkins, S. and J. Midgley. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association.* 35: .

Hughes, A., P. Trudgill and D. Watt. 2005. *English Accents and Dialects : an introduction to social and regional varieties of English in the British Isles*. Hodder Arnold.

Ingram, J. and S. Park. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics.* 25: 343 – 370.

Iverson, P. and P. K. Kuhl. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *The Journal of the Acoustical Society of America.* 97: 553-562.

Iverson, P., P. K. Kuhl, R. Akahane-Yamada, E. Dieschd, Y. i. Tohkura, A. Kettermann and C. Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition.* 87: B47–B57.

James M Scobbie, J. Stuart-Smith and a. E. Lawson, 2008. Looking variation and change in the mouth: developing the sociolinguistic potential of Ultrasound Tongue Imaging. Research Report for ESRC Project.

Jong, K. and B. Zawaydeh. 2002. Comparing Stress, Lexical Focus, and Segmental Focus: patterns of variation in Arabic vowel duration *Journal of Phonetics.* 30: 53-75.

Jong, K. d. 2004. Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics.* 32: 493–516.

Jongman, A., W. Herd and M. Al-Masri. 2007. A coustic Correlate of Emphasis in Arabic. *International Congress of Phonetic Sciences,* 913-916. Saarbrücken.

Joseph, M. and E. Odisho. 2005. *Techniques of Teaching Comparative Pronunciation in Arabic and English*. Gorgias Press LLC.

Khattab, G. 2007. A Phonetic Study of Gemination in Lebanese Arabic. In proceedings of the *Special session on 'Arabic at the beginning of the 2nd Millenium' at the the 16th International Congress of the Phonetic Sciences* 153-158. Saarbruecken, Germany: Universitaet des Saarlandes.

Khattab, G., F. Al-Tamimi and B. Heselwood. 2006. Acoustic and Auditory differences in the /t/- /T/ Opposition in Male and Female Speakers of Jordanian Arabic. In *Perspectives on Arabic Linguistics XVI: Papers from the sixteenth annual symposium on Arabic linguistics,* ed. Boudelaa, S., 131-160. Cambridge, UK: John Benjamins.

Khattab, G. and J. Al-Tamimi. 2008. Durational Cues for Gemination in Lebanese Arabic. *Language and Linguistics.* 22: 39-55.

Kiefte, M. and K. Kluender. 2005. The relative importance of spectral tilt in monophthongs and diphthongs. *Journal of the Acoustical Society of America.* 117: 1395–1404.

Klein, W. 1986. *Second Language Acquisition*. Cambridge: Cambridge University Press.

Kondaurova, M. V. and A. L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journa of the Acoustical Society of America.* 124: 3959–3971.

Kopczynski, A. and R. Meliani. 1993a. The Vowels of Arabic and English. *Papers and Studies in Contrastive Linguistics (PaSiCL).* 27: 183-192.

Kuhl, P. and P. Iverson. 1995. Linguistic experience and the "perceptual magnet effect". In *Speech perception and linguistic experience: Issues in cross-language research,* ed. Strange, W., 121 - 154. York Press: Baltimore, MD.

Kuhl, P. and J. Miller. 1975. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science.* 190: 69-72.

Kuhl, P. K. 1994. Learning and representation in speech and language. *Current Opinion in Neurobiology.* 4: 812–822.

Kuhl, P. K. 2000. A new view of language acquisition. In proceedings of the *National Academy of Sciences colloquium ‘‘Auditory Neuroscience: Development, Transduction, and Integration",* 11850–11857. Arnold and Mabel Beckman Center in Irvine, CA.

Kuhl, P. K., B. T. Conboy, S. Coffey-Corina, D. Padden, M. Rivera-Gaxiola and T. Nelson. 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society.* B: 979-1000.

Kuhl, P. K. and A. N. Meltzoff. 1996. Infant vocalizations in response to speech: vocal imitation and developmental change *Acoustic Society of America.* 100: 2425-2438.

Ladefoged, P. 1996. *Elements of Acoustic Phonetics*. University Of Chicago Press.

Ladefoged, P. 2006. *A Course in Phonetics*. Boston: Thomson Wadsworth.

Ladefoged, P. and D. E. Broadbent. 1957. Information Conveyed by Vowels. *The Journal of the Acoustical Society of America.* 29: 98-104.

Lado, R. 1957. *Linguistics Across Cultures*. University of Michigan Press, Ann Arbor.

Lengeris, A. and Hazan.V. 2007. Cross-language perceptual assimilation and discrimination of Southern British English vowels by Greek and Japanese learners of English. In proceedings of the *Proceedings of the 16th International Congress of Phonetic Sciences,* 1641-1644. Saarbrücken, Germany.

Liberman, A., F. Cooper, D. Shankweiler and M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review.* 74: 431-61.

Liberman, A., K. Harris, H. Hoffman and B. Griffith. 1957. The Discrimination of Speech Sounds Within and Across Phoneme Boundaries. *Journal of Experimental Psychology.* 54: 358–368.

Liberman, A. and I. Mattingly. 1985. The Motor Theory of Speech Perception Revised. *Cognition.* 21: 1-36.

Lisker, L. and A. Abramson. 1970. The voicing dimension: some experiments in comparative phonetics. In proceedings of the *6th Int. Congr. Phon. Sci.*, 563–67. Prague: Academia.

McMahon, A. 2002. *An Introduction to English Phonology*. Edinburgh University Press.

Mitchell, T. F. 1993. *Pronouncing Arabic 2*. Oxford: Clarendon Press.

Mitleb, F. 1981. Timing of English vowels spoken with an Arabic accent. *Research in Phonetics Report. Bloomington: Indiana University, Department of Linguistics* 193-226.

Mitleb, F. 1982. Voicing effect on vowel duration is not an absolute universal (A). *Journal of the International Phonetic Association.* 71: 23.

Mitleb, F. M. 1984b. Vowel length contrast in Arabic and English: A spectrographic test. *Journal of Phonetics.* 12: 229-35.

Mora, J. C. 2005. Lexical Knowledge Effects on the Discrimination of Non-Native Phonemic Contrasts in Words and Non-Words by Spanish/Catalan Bilingual Learners of English. In proceedings of the *ISCA Workshop on Plasticity in Speech Perception (PSP2005),* 43-46. London.

Munro, M. 1993. Productions of English Vowels by Native Speakers of Arabic: Acoustic Measurements and Accentedness Ratings. *Language and Speech.* 36: 39-66.

Nabelek, A., Z. Czyzewski and L. Krishnan. 1992. The Influence of Talker Differences on Vowel Identification by Normal-hearing and Hearing-imapaired Listeners. *Journal of the Acoustical Society of America.* 92: 1228-1246.

Newman, D. 2002. The phonetic status of Arabic within the worlds languages: the uniqueness of the lughat al-daad. *Antwerp Papers in Linguistics.* 100: 65-75.

Newman, D. and J. Verhoeven. 2002. Frequency analysis of Arabic vowels in Connected Speech. *Antwerp Papers in Linguistics.* 100: 77-86.

Niedzielski, N. 1999. The Effect of Social Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology.* 18: 62-85.

Paolo, M. D., M. Yaeger-Dror and A. B. Wassink. 2011. Analyzing vowels. In *Sociophonetics: A student's guide,* ed. Yaeger-Dror, M. D. P. a. M., 87-107. Routledge.

peterson, G., & Lehiste, I. 1961. Transitions, Glides, and Dipthongs. *Journal of the Acoustical Society of America.* 33: 268-277.

Pilus, Z. 2005. The perception of voicing in English word-final consonants: Examining the perceptual assimilation model. *Malaysian Journal of ELT Research.* 1: .

Pisoni, D. B. 1997. Some thoughts on "normalization" in speech perception. In *Talker Variability in Speech Processing,* ed. Johnson, K. and J. W. Mullennix, 9-32. San Diego: Academic.

Roach, P. 1991. *English Phonetics and Phonology*. Cambridge: Cambridge University Press.

Rosner, B. S. and J. B. Pickering. 1994. *Vowel Perception and Production*. Oxford Science Publications.

Ryding, K. C. 2005. *A reference Grammar of Modern Standard Arabic*. Cambridge University Press.

Scobbie, J. M. 2007. Biological and Social Grounding of Phonology: Vatiation as a Research Tool. *ICPhS.* XVI: 225-228.

Scobbie, J. M., J. Stuart-Smith and E. and Lawson, 2008. Looking variation and change in the mouth: developing the sociolinguistic potential of Ultrasound Tongue Imaging. Research Report for ESRC Project.

Smith, B. 2001. Arabic Speakers. In *Learner English: a teacher's guide to interference and other problems,* ed. Swan, M. and B. Smith, 195-214. Cambridge University Press.

Stevens, K. 1989. On the quantal nature of speech. *Journal of Phonetics.* 17: 3–45.

Stevens, K. and A. House. 1963. Perturbation of Vowel Articulations by Consonantal Context: an Acoustical Study. *Journal of Speech and Hearing Research.* 6: 111-128.

Stevens, K. N. 1972. The quantal nature of speech: Evidencefrom articulatory-acoustic data. In *Human communication: A unified view,* ed. Denes, P. B., & David, Jr E.E. , 51–66. NewYork: McGraw Hill.

Stevens, K. N. 2002. Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America.* 111: 1872-1891 .

Stevens, K. N. and S. J. Keyser. 2010. Quantal theory,enhancement and overlap *Journal of Phonetics.* 38: 10–19.

Strange, W. and V. L. Shafer. 2008. Speech perception in second language learners: the re-education of selective perception. In *Phonology and second language acquisition,* ed. Hansen Edwards, J. G., and Zampini, M. L. , 153– 92. Philadelphia: John Benjamins.

Strange, W., R. Yamada, R. Kubo, S. Trent, K. Nishi and J. Jenkins. 1998. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics.* 26: 311-344.

Thomas, E. R. and T. Kendall. 2007. NORM: The vowel normalization and plotting suite. Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/ 26-12-2010.

Tsukada, K., D. Birdsong, E. Bialystok, M. Mackd, H. Sung and J. Flege. 2005. A Developmental Study of English Vowel Production and Perception by Native Korean Adults and Children. *Journal of Phonetics.* 33: 263-290.

Watson, J. 2002. *The Phonology and Morphology of Arabic*. Oxford University Press.

Wells, J. C. 1986. *Accents of English 2: The British Isles*. Cambridge University Press.

Wood, D. 2004. An Empirical Investigation into the Facilitating Role of Automatized Lexical Phrases in Second Language Fluency Development. *Journal of Language and Learning.* 2: 1740 - 4983.

Yang, B. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics.* 24: 245 – 261.

Youssef, I. 2010. Against Underlying Mid Vowels in Cairene Arabic. *Zeitschrift für Arabische Linguistik (Journal of Arabic Linguistics). Wiesbaden: Harrassowitz.* 52: 5–38.

10 Appendix A: Language Background Questionnaire

10.1 Questionnaire for Syrian Arabic participants

Questions in red were designed for the participants who were in the UK and spent sometime in the UK.

The purpose of this questionnaire is to learn something about your language history. I would like to find out what languages you know, when you first learned them, and how much you use them. I also would like to know about your English language formal learning in Syria.

Today's date: ________

1. Your name: ____
2. Gender:____
3. City and province of birth: ______________________________
4. Date of birth: ______________________________________
5. What is your first 1st language? ________ 2nd ________3rd ________
6. What is the city or place of birth of your parents?

   Father ____

   Mother ____
7. How do you classify your English input (British English or American English)?
8. Your age of arrival in United Kingdom:

   ____
9. Years you have lived in United Kingdom:

   ____
10. Years and places you have lived around in the United Kingdom:

	Place or city in the UK	Years (Length of residence)
1		
2		
3		
4		

11. Please estimate, using a percentage (%) how much you have spoken Syrian Arabic in the past 5 years _____; in the past 5 months _____; in the past 5 weeks _____.
12. Number of years you have spent in an English speaking country (if any)

____

13. number of years of English formal education in Syria (including school and university) _______ and in UK _______.
14. Please estimate your ability to speak, understand, read, and write English. Use the number "1" if your ability is poor, "7" if your ability is good, and numbers in between for ability levels that are in between.

	English						
	1	2	3	4	5	6	7
Speaking							
Understanding							
Reading							
Writing							

15. Please, state if there are any *speech or hearing disorders*

____

10.2 Questionnaire for Native English participants

The purpose of this questionnaire is to learn something about your language history. I would like to find out what languages you know, when you first learned them, and how much you use them.

Today's date: ________

1. Your name: ____
2. Gender:____
3. City and province of birth: __________________________________
4. Date of birth: ____
5. What is your first 1st language? ______________________________
6. Languages you have studied:

	Language studied	Years of formal education
1		
2		
3		
4		

7. Please state if you master other languages in a native-like way or as a bilingual speaker:

   ____

8. Years and places you have lived around in the United Kingdom:

	Place or city in the UK	Years (Length of residence)
1		
2		
3		
4		

9. Number of years you have spent in any other country (if any)

   ____

10. Please, state if there are any *speech or hearing disorders.* ____________

11 Appendix B: Identification Task English alternative responses

	First choice		Second choice		Third choice	
1	a.	Pot	b.	Put	c.	Poot
2	a.	Sheet	b.	Shoot	c.	Shirt
3	a.	Bays	b.	Buys	c.	Boys
4	a.	Dears	b.	Dies	c.	Dares
5	a.	Bet	b.	Bit	c.	Beat
6	a.	Curt	b.	Cart	c.	Cut
7	a.	Bit	b.	Bat	c.	Bet
8	a.	Toon	b.	Town	c.	Tone
9	a.	Bird	b.	Bud	c.	Bard
10	a.	Bed	b.	Bard	c.	Bud
11	a.	Beat	b.	Bit	c.	Bet
12	a.	Beat	b.	Bat	c.	Bet
13	a.	Fears	b.	Phase	c.	Fairs
14	a.	Bard	b.	Bud	c.	Bed
15	a.	Bees	b.	Boys	c.	Buys
16	a.	Had	b.	Hood	c.	Hard
17	a.	Bought	b.	Boot	c.	Put
18	a.	Curt	b.	Coat	c.	Cart
19	a.	Fairs	b.	Fears	c.	Phase
20	a.	Bit	b.	Boot	c.	Baught
21	a.	Shoot	b.	Short	c.	Shot
22	a.	Tone	b.	Town	c.	Toon
23	a.	Cot	b.	Cart	c.	Cut
24	a.	Days	b.	Dears	c.	Dies
25	a.	Port	b.	Put	c.	Pot
26	a.	Known	b.	Noun	c.	Nine
27	a.	Shut	b.	Sheet	c.	Shot
28	a.	Heat	b.	Hat	c.	Hot
29	a.	Phase	b.	Fares	c.	Fees
30	a.	Dies	b.	Dares	c.	Days
31	a.	Port	b.	Put	c.	Pot
32	a.	took	b.	turk	c.	tick
33	a.	Doers	b.	Dares	c.	Dears
34	a.	Bed	b.	Bud	c.	Bird
35	a.	Cut	b.	Cat	c.	Cot
36	a.	Fight	b.	Feet	c.	Fate
37	a.	Poot	b.	Pot	c.	Put
38	a.	Cut	b.	Cart	c.	Cat

39	a.	Bays	b.	Boys	c.	Buys
40	a.	Nine	b.	Noun	c.	Known
41	a.	Bud	b.	Bed	c.	Bard
42	a.	Bert	b.	Bet	c.	Bought
43	a.	Cut	b.	Cot	c.	Cart
44	a.	Dears	b.	Dares	c.	Doers
45	a.	Short	b.	Shut	c.	Shoot
46	a.	turk	b.	took	c.	tick
47	a.	Phase	b.	Fairs	c.	Fears
48	a.	Hard	b.	head	c.	Hood
49	a.	Shot	b.	Short	c.	Shoot
50	a.	Cut	b.	Cart	c.	Cat
51	a.	Bet	b.	Bought	c.	Bert
52	a.	Port	b.	Put	c.	Pot
53	a.	Days	b.	Dears	c.	Dares
54	a.	Hard	b.	Hide	c.	Howdy
55	a.	Bit	b.	Beat	c.	Bet
56	a.	Heart	b.	Hot	c.	Hat
57	a.	Sheet	b.	Shoot	c.	Shirt
58	a.	But	b.	Bought	c.	Boot
59	a.	Feet	b.	Fight	c.	Fate
60	a.	Toys	b.	Tors	c.	Tours
61	a.	Pot	b.	Put	c.	Poot
62	a.	tick	b.	took	c.	turk
63	a.	Town	b.	toon	c.	Tone
64	a.	Shot	b.	Shut	c.	Shirt
65	a.	Bat	b.	Bet	c.	Bit
66	a.	Cot	b.	Cut	c.	Cart
67	a.	Put	b.	Boot	c.	Bought
68	a.	Hear	b.	Her	c.	Hair
69	a.	Noun	b.	Known	c.	Nine
70	a.	Bat	b.	Put	c.	Pot
71	a.	Cat	b.	Cut	c.	Caught
72	a.	Toy	b.	Tour	c.	Tore
73	a.	Boys	b.	Buys	c.	Bays
74	a.	Cat	b.	Cot	c.	Cut
75	a.	Far	b.	Fear	c.	Fare
76	a.	Bet	b.	Bit	c.	Beat
77	a.	Shoot	b.	Shirt	c.	Sheet
78	a.	Poot	b.	Pot	c.	Put
79	a.	Curt	b.	Cut	c.	Cart
80	a.	Pot	b.	Port	c.	Put
81	a.	Shot	b.	Shoot	c.	Short
82	a.	Boat	b.	Bought	c.	Bout

83	a.	Bud	b.	Bard	c.	Bird
84	a.	Her	b.	Hear	c.	Hair
85	a.	Boot	b.	But	c.	Bat
86	a.	Bit	b.	Bat	c.	Bet
87	a.	Toon	b.	Town	c.	Tone
88	a.	Beat	b.	Bet	c.	Bit
89	a.	Fees	b.	Fears	c.	Phase
90	a.	But	b.	Pot	c.	Bet
91	a.	Put	b.	Pot	c.	Port
92	a.	Beat	b.	Bat	c.	Bet
93	a.	Hide	b.	Howdy	c.	Hard
94	a.	Bed	b.	Bird	c.	Bud
95	a.	Bat	b.	Bit	c.	Beat
96	a.	Bought	b.	But	c.	Boot
97	a.	Fears	b.	Fees	c.	Phase
98	a.	Bat	b.	Bet	c.	Bit
99	a.	Cut	b.	Curt	c.	Cart
100	a.	Hat	b.	Hot	c.	Heart
101	a.	Bat	b.	Beat	c.	Bit
102	a.	Toes	b.	Tears (paper)	c.	Towers
103	a.	Bought	b.	Put	c.	Boot
104	a.	Bird	b.	Bed	c.	Bud
105	a.	Bought	b.	Bert	c.	Bet
106	a.	Cart	b.	Cat	c.	Cut
107	a.	Beat	b.	Bet	c.	Bit
108	a.	Bet	b.	Bit	c.	Bat
109	a.	Poot	b.	Pot	c.	Put
110	a.	Tears (paper)	b.	Towers	c.	Toys
111	a.	Pot	b.	Port	c.	Put
112	a.	Cot	b.	Cut	c.	Cat
113	a.	Beat	b.	Bit	c.	Bat
114	a.	Bud	b.	Bard	c.	Bird
115	a.	Bit	b.	Bet	c.	Bat
116	a.	Poot	b.	Put	c.	Pot
117	a.	Tool	b.	Toy	c.	Tour
118	a.	Hot	b.	Heart	c.	Hat
119	a.	Bet	b.	Beat	c.	Bat
120	a.	Towers	b.	Tears (paper)	c.	Toys