Doctoral (PhD) Dissertation

The Perception and Production of American English Sounds by Palestinian Arabic Adolescents

DOI:10.18136/PE.2022.820

By

Bashar M. M. Farran

Supervisors:

Prof. Dr. Vincent J. van Heuven

Dr. Ildikó Hortobágyi

Multilingualism Doctoral School

Faculty of Modern Philology and Social Sciences

University of Pannonia

Veszprém, 2022

STATEMENT

This dissertation, written under the direction of the candidate's dissertation committee and approved by the members of the committee, has been presented to and accepted by the Faculty of Modern Philology and Social Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The content and research methodologies presented in this work represent the work of the candidate alone.

Bashar M. M. Farran	, 2022
Candidate	Date

Dissertation Committee:

________________________________	, 2022
Chairperson	Date
________________________________	, 2022
First Reader	Date
________________________________	, 2022
Second Reader	Date

The Perception and Production of American English Sounds by Palestinian Arabic Adolescents

Thesis for obtaining a PhD degree in the Doctoral School of Multilingualism of the University of Pannonia

in the branch of Applied Linguistics

Written by Bashar M. M. Farran

Supervisors: Prof. Dr. Vincent J. van Heuven

Dr. Ildikó Hortobágyi

Propose acceptance (yes / no) …………….…………….

(supervisor/s)

As a reviewer, I propose acceptance of the thesis:

Name of Reviewer: …………………… yes / no

…..…………….…………….

(reviewer)

Name of Reviewer: …………………… yes / no

……...………….…………….

(reviewer)

The PhD-candidate has achieved …………% at the public discussion.

Veszprém, ……/… / 2022 …………….…………….

(Chairman of the Committee)

The grade of the PhD Diploma …..................................... (……. %)

Veszprém, ……/… / 2022 …………….…………….

(Chairman of UDHC)

ABSTRACT

This dissertation addresses the perception and production of American English monophthongs by Palestinian Arabic (PA) learners of English as a Foreign Language (EFL). It aims to identify (and predict) areas of difficulty in the perception and production of these sounds of American English (AE) within a framework of the most influential L2 perception and production theories and models and their most recent versions so that teaching (materials) can address these rather than spend time on sounds that do not constitute a problem. By devising three separate yet interrelated studies, PA learners' perception and production of AE as EFL was explored and compared with similar data collected from native AE participants (van Heuven et al., 2020; Wang & Van Heuven, 2006).

The first study investigated PA learners' perceptual assimilation of AE vowels. The study examined how the eleven AE monophthongs (Ladefoged, 1999: 41) map onto the six vowels of PA (Thelwall & Sa'adeddin, l999: 52). The Perceptual Assimilation Model (PAM) predicts learning problems when two L2 phonemes are perceived as equally good tokens of a Single Category in the L1 (SC scenario) (Best, 1995). SC contrasts will yield an incorrect perceptual representation of the AE vowel system in the mind of the (beginning) PA EFL learner, with insufficient spectral or temporal separation of categories (compared to native AE listeners). Forty (20 male and 20 female), adolescent PA high-school learners of EFL listened to the monophthongs of AE (four tokens of each, in different random orders per participant) spoken in /hVd/ words, and classified these as one of the six PA vowels /i, i:, a, a:, u, u:/ while rating them on a 5-point goodness scale. Seven SC contrasts were identified in the results, i.e., *heed-hayed* /i:-e:/, *hid-head* /ɪ-ɛ/, *hud-hood* /ʌ-ʊ/, *hod-hawed* /ɑ:-ɔ:/, *hawed-hoed* /ɔ:-o:/, *hawed-who'd* /ɔ:-u:/, and *hoed-who'd* /o:-u:/. Moreover, a Category Goodness (CG, intermediate difficulty predicted) problem was identified for the *had-hod* /æ:-ɑ:/ contrast. Contrasts that rely on a difference in vowel length did not cause any problems. The results of this study showed that there is a general confusion in mapping AE vowels within the same length category if they are spectrally close (i.e., have similar vowel quality). A comparison was made for the PA EFL results with results from similar studies on other L1 Arabic varieties' perception of AE and revealed conformity between PA perception of AE vowels and the other studies' results in the sense that the L1 Arabic perception of AE is not exclusively differentiated based on duration but also includes spectral differences between long and short counterparts. The results of the comparison also showed differences in relation to which AE vowels were the most confused

ones. Finally, the results of the first study provided predictions and hypotheses for the second study concerning the mental representation of the AE vowel space in the minds of PA learners.

In the second study, the same 40 PA participants listened to and identified 86 artificial vowel sounds (7 degrees of height, 9 degrees of backness/rounding, and 2 lengths) sampled with perceptually equal steps along the F1 and F2 dimensions of the vowel space, excluding 20 impossible combinations (Van Heuven et al. 2020), in /mVf/ nonwords with vowel durations of 200 or 300 ms. Listeners identified each token as one of the eleven AE monophthongs while rating them on a 3-point goodness scale. The experiment was repeated with a control group of 20 (10 male and 10 female) native AE listeners. This study aimed to reveal the differences in the mental representation of the vowel space between native AE listeners and nonnative PA learners of AE. The main objective of this study was to determine whether nonnative listeners perceive the AE vowel space the same way as native AE listeners do and, if not, what their perceptual representation looks like in terms of vowel quality as determined by formant structure, duration, and the relative importance (trading relationship) between quality and duration. The results show that the PA learners' conception of the AE vowel system is incorrect in several important respects. The PA participants' perceptual representation of the AE vowels differed from the native AE norm and was strongly influenced by the vowel system of PA. Vowel duration proved a much more important characteristic for PA listeners than for AE controls, as EFL learners relied almost exclusively on vowel duration to differentiate spectrally adjacent vowels (in *feel-fill* or *fool-full*), while native listeners of AE relied on vowel quality rather than length (confirming Hillenbrand et al., 2000). Additionally, the EFL learners accepted monophthongal /e:/ and /o:/ (as in *sale* and *whole*), which were rejected by the native listeners because of insufficient diphthongization. The vowels in *fill-tell* were not differentiated, and most mid-low vowel sounds were incorrectly identified as /ʌ/ (as in *null*). This study concluded that these confusions require serious attention at the pedagogical level, as they will most likely lead to pronunciation errors. The argument is that the structure of the PA vowel system, with its three-point vowels /i, a, u/ and a phonemic length contrast (short, long), is the underlying cause of the flawed perceptual representation of the AE vowels, which, in turn, is hypothesized to cause deviations from the AE norms in the PA participants' speech production.

Based on the predictions of the previous two studies, the third study aimed to assess PA learners' production of AE monophthongs through a qualitative and quantitative analysis to provide a comprehensive description of AE vowels produced by PA learners. The study measured the articulation of the 11 monophthongs of AE in /hVd/ words (e.g., *heed*, *hid*, *head*, *had*, …) by the 40 PA learners of EFL and compared their results with results of 20 (10 male and 10 female) native AE university students for the same test (Wang & Van Heuven, 2006). In general, the results of this chapter show a clear carry-over of Arabic phonetic spectral attributes. The learners' results showed a great deal of overlap in producing AE new vowels. Only vowels with unique counterparts (i.e., AE high front/back long /i/, /æ/, and /u/) between the two vowel inventories showed a distinct distribution from their spectrally adjacent competitors. All other AE vowels were confused in three vowel clusters in a partially overlapping manner. Vowel durations of the L2 and L1 speakers proved strongly correlated but were (much) shorter in L2 than in L1. To conclude this study, the nonnative PA production results show confused alignment of the L2 AE vowel system, especially for the new AE vowels.

The learner's L1 affects the production of EFL vowels in general, yielding a foreign accent that diverges from the native English norms and resembles EFL learners' L1 more. Therefore, the PA results in the production study were compared with those found in the literature for EFL learners with Arabic L1 backgrounds other than PA, both spectrally and temporally, to inspect whether their different L1 Arabic varieties affect their production of AE differently. The results show both differences and similarities between the speakers of what is often considered one shared L1, i.e., Arabic. Among the similarities are the overall shrinking of the AE L2 vowel space (relative to L1 control data in Wang & Van Heuven, 2006) and the systematic overshortening of all vowel durations. The remaining differences in the spectral organization of the nonnative vowel spaces can be attributed to differences in the L1 varieties. Depending on how the differences in the AE vowels impact the nonnatives' intelligibility, this dissertation provided some pedagogical implications and recommendations for EFL curriculum developers, teachers, and learners that require different tailer-cut remedies for each regional variety of Arabic.

ACKNOWLEDGEMENT

I would like to thank the chair of the committee and the dean of the doctoral school, Prof. Dr. Judit Navracsics for her professional guidance, advice, and all efforts during the course of my research and PhD study.

I am grateful to Prof. Dr. Vincent van Heuven for his professional guidance, help, and support in training me in linguistic fieldwork and experimental linguistics. This project would not have been possible without his professional advice and generous time in revising the numerous drafts of this dissertation. I am eternally grateful to have been able to carry out this work under his supervision.

I am deeply indebted to Dr. Ildikó Hortobágyi for her invaluable guidance, patience, and encouraging me to always think critically. I am proud of, and grateful for, my time working with her.

Additionally, I would like to express my gratitude to the members and opponents of my dissertation committee, i.e., outside examiners Prof. Dr. Ferenc Bunta of the University of Houston (Texas, USA), Dr. Tekla Gráczi of the Hungarian Academy of Sciences (Budapest), and internal opponent Dr. Szilárd Szentgyörgyi. Also, I would like to personally thank and express my gratitude to Dr. Szilvia Bátyi, who has never hesitated in helping students with her academic and professional knowledge.

Thanks to the University of Pannonia and its Multilingualism Doctoral School for everything they provided for me. I would like also to thank Stipendium Hungaricum for awarding me a PhD scholarship for 4 years, and the Hungarian Ministry for Innovation and Technology for awarding me the Únkp-21-3 New National Excellence Fellowship, providing me with the financial means in the final year of my PhD journey to complete this project.

I would like to thank the Ministry of Education in Palestine for facilitating my field research in the high school in the West Bank. Additionally, all participants in this study deserve a note of appreciation and thanks, for volunteering their time and efforts.

Additionally, I would like to praise my PhD friends and colleagues (Faten Amer, Anna Ismail, Ibtisam Smari, Rania Salah, and Aladdin Khalifa) whom I had the opportunity to exchange ideas and knowledge. I'm also indebted to my colleague and great friend Aya M. Halabi, Dr-to-be at the University of Michigan-Ann Arbor for always encouraging and supporting me with persistent strength. I wish you a great future ahead, full of passion, love, and great success.

Finally, I am most thankful to my wonderful family for their patience and support to finish this study, especially my father, Mohammad Ata, and my brothers Amer and Ammar for all their sacrifices and unconditional love. Last but not least, I'm especially indebted to Nariman Salameh who always supported and encouraged me with her words and actions to do the best and keep going. I wish you eternal happiness and a great future.

DEDICATION

To my mother, Huda

My first teacher, who despite her limited knowledge of English, taught me the alphabet of the language and always encouraged me to be the best version of myself.

May Allah (SWT) heal and bless you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AE	American English
CA	Classical Arabic
CAH	Contrastive Analysis Hypothesis
CAV	Colloquial Arabic Variety
CC	Consonant Cluster
CG	Category Goodness
CU	Categorized Uncategorized
CV	Consonant Vowel
EFL	English as a Foreign language
F0	Fundamental Frequency
F1	First Formant
F2	Second Formant
F3	Third Formant
GAE	General American English
ISCH	Interlanguage Structural Conformity Hypothesis
L1	First Language
L2	Second Language
L2LP	Second Language Linguistic Perception
L3	Third Language
LA	Levantine Arabic
LBQ	Language Background Questionnaire
LP	Linguistic Perception
LPC	Linear Predictive Coding
MDH	Markedness Differential Hypothesis
MENA	Middle East and North Africa region
MFC	Multiple Forced Choice
MoE	Ministry of Education
MSA	Modern Standard Arabic
OT	Optimality Theory
PA	Palestinian Arabic
PAM	Perceptual Assimilation Model
SC	Single Category
SSBE	Standard Southern British English
SLA	Second Language Acquisition
SLM	Speech Learning Model
TC	Two-Categories
TL	Target Language

CHAPTER ONE

INTRODUCTION

1.1. Introduction

The present thesis investigates the perception and production of American English (AE) segmental features (sounds) by a group of Palestinian-Arabic (PA) adolescent learners of English through a well-founded theoretical approach.[1] More particularly, the present research sheds light on the current areas of difficulty that PA learners face with the perception and production of AE vowel sounds through three separate but integrated research studies. Study 1 tackles the perception of the AE pure vowels by the PA learners through an assimilation task to the learners' L1 vowel inventory to identify (and predict) those target vowels that may constitute learning problems for L2 learners. Study 2 addresses the PA learners' perceptual representation of the AE pure vowels through an identification task that consists of 86 synthesized vowels to be identified as one of the AE pure vowels. Study 3 consists of a production task where learners were requested to produce three tokens of words that contain the target monophthongs. The reported results of Study 1 were compared with the results of some studies on varieties of Arabic other than PA. The results of studies 2 & 3 (perception and production data) were compared with similar data collected from native AE listeners/speakers (Van Heuven et al., 2020; Wang & Van Heuven, 2006), and the results of study 3 on vowel production were also compared to varieties of Arabic other than PA. My predictions are based on partially competing principles of several recent L2 and cross-linguistic theories: the Contrastive analysis hypothesis (CAH, Lado, 1957; Wardhaugh, 1970), Markedness Differential Hypothesis (MDH, Eckman, 1977), Interlanguage Structural Conformity Hypothesis (ISCH, Eckman, 1991), Perceptual Assimilation Model (PAM, Best, 1994; 1995; Best et al., 2001) and its extension to L2 learning (PAM-L2, Best & Tyler, 2007; Tyler, 2019), Speech Learning Model (SLM, Flege, 1995, 2002) and its revised version (SLM-r, Flege & Bohn, 2021), and Second Language Linguistic Perception (L2LP, Escudero, 2005; 2009) and its revised version (Van Leussen & Escudero, 2015). These models and theories are employed to account for the different studies performed in this thesis and their results.

This chapter starts by presenting an overview of the English language reality in Palestine in section 1.2 and from there to the MENA region situation in section 1.3 while highlighting the

[1] English in this dissertation refers to the American English unless otherwise specified.

diglossia status with a modern lens, which leads to the purpose of the current thesis in section 1.4. The research questions and hypotheses are presented in section 1.5. The penultimate part of this chapter is section 1.6, where the importance of this study is presented, and finally, section 1.7 outlines the overall structure of the dissertation.

1.2. Overview of English reality in Palestine

Palestinian Arabic (PA) belongs to the Levantine variety of Arabic, which is spoken along the coast of the Levantine Sea. It includes speakers from Palestine, Jordan, Syria, Lebanon, Turkey, and Arabs inside Historical Palestine. The PA variety is spoken by over 8.5 million people (Shahin, 2011). Among the speakers of PA, approximately 3 million reside in Jordan, where they contributed substantially to shaping the dialects of the Jordanian kingdom, especially the urban one (Al-Wer, 2002, 2007; Herin, 2013). Additionally, the Arab community in Israel is approximately 2 million people of those who have remained in historical Palestine after it came under the control of Israel after the year 1948 and again in 1967 (Horesh, 2020). Due to the Palestinian diaspora, many speakers of PA are now living elsewhere in the Middle East and around the world, especially in Latin America.

Palestine is a country with a vast and diverse history of linguistic interaction. Amara (2003) reported that more than 3 languages acted as the official languages in the area that is now known as Palestine during the last 200 years, as Turkish, English, and Hebrew were (or still are) spoken as official languages along with the local and original language, i.e., Arabic. These languages descend from different families. Despite the many occupations that Palestine has witnessed and still does, indigenous Palestinians are still committed to their Arabic language because of its identificatory, sociological, and religious factors. While languages other than Arabic still exist as minority or liturgical languages, English has a foreign language (EFL) status, as it is taught at schools for 12 years and is vastly present in the linguistic landscape of the country (Farran & Hortobágyi, 2020).

Developing a nation's language education requires continuous analysis of the content and the results of language examinations. The current picture of English language education in Palestine – including the West Bank, East Jerusalem, and the Gaza Strip – reflects high levels of literacy, up to 98%, according to the 2018 United Nations and World Bank reports. In Palestine, students start officially learning English in the first grade, so all PA L1 learners of English start learning English as a foreign language at the age of five. The formal language tests as part of the EFL output assessment, especially in high school leaving exams, do not fully assess the four language skills. Especially the speaking and listening skills are incompletely tested (Farran et al., 2020).

One of the most difficult aspects of learning a second or foreign language (L2) is acquiring the correct sounds of it. Such difficulty results in failing to achieve a native-like level of performance, which is prompted by many linguistic and social factors (e.g., different sound inventories, different orthographies, stress patterns, etc.). The salient result of this difficulty is obvious and characterizes adult L2 learners' perception and production of what is generally called "foreign accent", which may impede communication and contribute to social stigma and discrimination (Gluszek & Dovidio, 2010; Hosoda & Stone-Romero, 2010; Collins & Clément, 2012; Buckingham, 2014). In alignment with the formal objective of teaching English in Palestine, speaking and listening skills fall at the heart of the dilemma (Farran et al., 2020).

Studies on L2 English perception and production by L1 Arabic learners are few and far between. Relatively recently, studies have begun to tackle the acquisition and production of English by certain L1 Arabic varieties. Nevertheless, interference by L1 Palestinian Arabic (PA) has not been studied in depth before now.

1.3. Overview of Arabic language and diglossia

One of the pioneering researchers of phonological variation who studied Arabic dialects is Charles A. Ferguson. In a study conducted in (1959), he distinguished between two varieties of the language, namely, High Arabic and Low Arabic. These two different forms of Arabic exist in the society side by side: the formalized Modern Standard Arabic (MSA) and the Colloquial Arabic varieties (CAVs). The pan-Arab MSA corresponds to the High Variety, which is used as a vehicle for "Highly Codified" literature and liturgical purposes and is learnt at schools. Nevertheless, colloquial Arabic corresponds to the Low Variety that is used in everyday casual speech. Despite being deeply entrenched in MSA, colloquial dialects differ considerably at the articulation (pronunciation), syntax, and semantic levels. Ferguson (1959) also showed that the High Variety is considered superior to the Low Variety, as it is more prestigious and is related to religious ends, which justifies why most studies focus on this variety in comparison to other CAVs. However, these stratifications of Arabic vernacular varieties are not as simple and dichotomous as suggested by Ferguson (1959); rather, CAVs are grouped into at least five, sometimes six, major dialect clusters based on geography and linguistic features. These varieties are Gulf Arabic (GA), Iraqi Arabic (IA), Maghrebi Arabic (MA), Yemeni Arabic (YA), Egyptian Arabic (EA), and Levantine Arabic (LA), which includes Palestinian Arabic

(Holes, 2004; Versteegh, 2014). These varieties differ on many linguistic levels, and their mutual intelligibility varies. See the map in Figure 1.1 below.

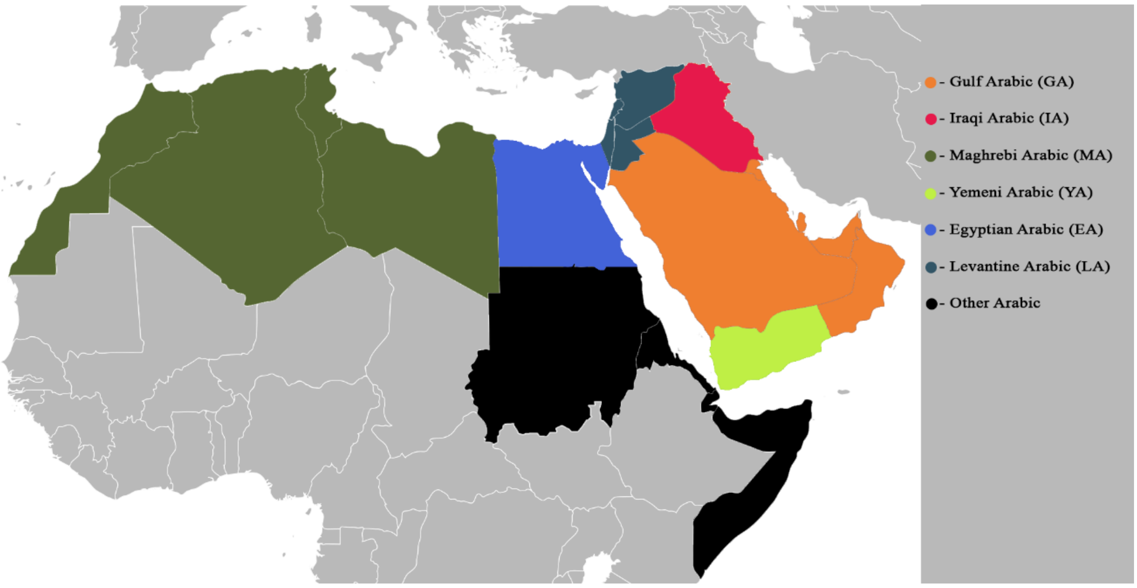

Figure 1.1. MENA region map, color-coded based on different spoken L1 Arabic varieties. Based on (Holes, 2004; Mustafawi, 2018; Owens, 2013: 19)

It is not straightforward to exactly predict what difficulties Arabic learners might face while acquiring English as L2, since the relationship between Arabic and English is complicated, especially at the phonetic and phonological levels. Furthermore, discussing the historical evolution of diglossia in the Middle East and North Africa (MENA) region is beyond the scope of this thesis. However, I would refer to the diglossic situation of Arabic as one of the main sources of these learning difficulties because there exist many variations between the low varieties of Arabic across the MENA region, even though they have a lot in common at the semantic and phonological levels, with MSA being the source for all of them. Therefore, the difficulty in acquiring a nonnative phone may well depend on the speaker's L1 low-variety of Arabic.

In addition to its diglossic dimension, Arabic has many varieties with specific features that are totally different from each other, especially at the phonotactic and phonological levels, which in return will have different effects on EFL acquisition. Several studies have analyzed EFL learners' perception and production acoustically while attempting to uncover the effects of L1 Arabic on learners' EFL proficiency. These studies have shed light on many aspects and levels of L2 sound acquisition and production. However, researchers have predominantly limited their perspective to the MSA phonological inventory as a general and acceptable representative of the language, rather than on the effects of different dialects or CAVs on EFL (Al-Badawi & Salim, 2014). To a lesser degree, other studies have focused on the effects of certain varieties of colloquial Arabic on the perception and/or production of EFL, e.g., Evans

& Alshangiti (2018) on the influence of Saudi Arabic; Faris et al. (2016) about Egyptian Arabic speakers of Australian English; Ali (2011) on Sudanese-accented British English (BE), among other studies that will be discussed in the following chapters.

1.4. Purpose of the study

At the phonological level, it is very difficult to predict or generalize the difficulties that L1 Arabic learners of English might encounter when acquiring the vowels and consonants of EFL unless the acquisition process of L2 for each L1 variety is inspected. This is because Arabic is spread across different countries and covers a wide geographical area. The differences between any two varieties of Arabic go beyond the syntax and the lexicon and involve the phonological level of the language.

Research on L2 sound acquisition found it challenging for L2 learners to acquire all new language nonnative sounds due to different factors, especially at the linguistic level and more particularly when the two languages have similar yet crucially differing phonetic sound categories. This problem with nonnative sound categorization might be attributed to an inability to properly perceive, and consequently produce, some phonetic categories of the target language (e.g., Bohn & Munro, 2007).

Since early phonological theories, the difficulties of producing L2 sounds are ascribed to the perception of speech sounds that are aligned to the L1 sounds and cannot be easily perceptually separated. Learners' categorical perceptual capabilities are calibrated to and developed by their L1 and its 'parameters'. Any non-L1 sound will fall within the already available identical L1 category, if available, or will cause a learning problem at the beginning in case it deviates from any L1 sound or a lasting learning problem in other cases. Trubetzkoy (1935 [1969]) encapsulated the situation as a 'phonological sieve' of the mother tongue that analyzes what is being heard within the parameters of L1, which yields many L2 mistakes and misinterpretations.[2] Based on this understanding, the majority of the current L2 theories and models are formulated.

Accordingly, investigating L2 sounds (in vowel perception and production) should pave the way to spot the interference effect of L1 on L2 and help to delineate the L1 PA sound characteristics and their effects on L2 AE sound perception and production.

To the best of my knowledge, no previous studies have addressed the perception of AE vowels by PA learners or have analyzed the difficulties that PA learners encounter when

[2] I cite here the English translation of Trubetskoy's book by Christina Baltaxe, which was originally published in German as *Grundzüge der Phonologie* in 1939.

learning to perceive AE vowels following an acoustic approach. The first study (Chapter 3) therefore attempts to fill the gap by addressing PA learners' perception of the AE vowels within the framework of the perceptual assimilation model (PAM & PAM-L2; Best, 1994; 1995; Best et al., 2001; Best & Tyler, 2007).

The main objective of the second study (Chapter 4) is to reveal the similarities and differences between the native AE and nonnative PA perception of the AE vowel space. Additionally, it aims to reveal how their mental and perceptual representations of the AE vowel space differ in terms of vowel quality (as determined by formants structure), quantity, and the relative importance (or trading relationship) between them.

In addition, no earlier studies have acoustically tested the PA variety production of English sounds in general and of AE in particular. Therefore, study 3 (Chapter 5) tackles the Palestinian Arabic production of the AE monophthongs. More data concerning the AE diphthongs and r-colored vowels have been collected but will not be included in this thesis. This was decided to maintain the homogeneity of comparison between the two vowel perception tests and the vowel production test. In addition, at the vowel level, it has been generally discussed that diphthongs are more intelligible than monophthongs, for instance, in English sung words (Johnson et al., 2014). In relation to this study, AE diphthongs (especially true diphthongs) are reportedly not known to create a pronunciation problem for Arabic learners of EFL (see, e.g., Rehman et al., 2020: 18). However, monophthongs are much more a source of reduced intelligibility.

The data available for other Arabic varieties from different previous studies as well as the results of the present research are anticipated to help L1 Arabic learners of English (in understanding the specificities of their L1 variety and how these might affect their L2 English acquisition); EFL teachers (in following different, and perhaps new, techniques while teaching sounds of AE), and curriculum developers in modifying, developing, and sifting the curriculum. It is expected that this research in particular will provide help with these matters for learners of AE in Palestine.

1.5. Research questions

The research studies conducted in this thesis try to answer the overall research questions of how properties of the sound system of a learner's native language may influence the acquisition of a second (foreign) language in secondary education. To this end, I conducted several studies, and each study answers its specific questions within this general frame. The answers accumulate to give a better understanding of the issues approached and contribute to my overall understanding of the acquisition of a nonnative sound system.

Study 1

1- How is the PA vowel system predicted to affect the perception of AE pure vowels?

i. Which AE vowels constitute a (potential) perception problem for PA EFL learners?

ii. (a) Which AE vowels are predicted to be the most difficult for PA learners to perceive correctly? and (b) How perceptually sensitive are PA learners to AE vowel duration and/or quality?

iii. How do the results of this research align with the results of similar studies on other CAVs learners of EFL?

Study 2

2- How perceptually sensitive are PA learners toward AE vowels' quantitative and qualitative features? And what are the most difficult AE vowels to perceive?

i. How do Palestinian Arabic listeners conceive of the American English vowel space?

ii. How do American native listeners conceive of their own vowel space?

iii. How does the mental representation of vowel sounds differ between AE native listeners and PA learners in terms of vowel quality and duration, and how do the quality and the duration interact or trade?

Study 3

3- How native-like is the production of AE monophthongs by PA learners in terms of duration and quality?

i. In terms of acoustic measurements of formants and durations, how do Palestinian Arabic speakers produce the American English vowels compared to native AE speakers?

ii. Which AE vowel contrasts are difficult to produce by PA listeners?

Two follow-up questions seemed necessary to be answered in light of the answers to the questions yielded in the third study. Namely,

iii. Is there a connection between PA learners' perception errors and their production (how does their perception affect their production of AE monophthongs)? and

iv. Do the PA results differ from nonnative Arabic L1 varieties other than PA with regard to the first two formants (F1, F2) and duration?

1.6. Hypotheses

I hypothesize, in general, that the L1 PA variety affects both the perception and the production of AE vowels and consonants due to mismatches between phonological features in L1 and L2. Each study (chapters 3-5) will contain more specific hypotheses that are justified within the context of the topic addressed.

1.7. Significance of the study

This study will be the first to investigate PA variety in depth and its possible phonological interference with L2 AE through an acoustic and experimental approach. My aim is to discover the L2 AE sound representation in the PA participants' minds and whether there is any process of merging or dividing the already existing sound categories to compensate for the mismatch with the AE sound inventory for consonants, consonant clusters, and especially for vowels, which immensely differ in numbers between the two approached languages. This thesis sets out with the aim to experimentally/acoustically assess the perception and production of AE sounds by L1 PA learners and to propose solutions for policy makers and teaching planners to overcome learners' accent problems.

In addition, since this thesis initiates a better understanding of EFL proficiency for Arabic L1 participants in Palestine, it contributes to the growing body of research on L1 Arabic varieties and their different effects on English as L2 acquisition and production.

1.8. Outline of the dissertation

Following this succinct overview of the context of the research, general research questions, the hypotheses, etc. The remainder of the thesis is organized as follows. The first part of Chapter 2 provides a contrastive analysis between MSA as a standard L1, PA as a spoken regional L1 variety, and AE as L2 for PA participants, while the other part of Chapter 2 reviews and discusses L2 speech perception and production theories, models of L2 acquisition, and related hypotheses. In Chapter 3 (Study 1), I investigate the perceptual assimilation of the 11 AE monophthongs to the PA vowel inventory by Palestinian Arabic (PA) adolescents who learn English as a Foreign Language (EFL) in Palestine. Next, in chapter 4 (Study 2), I provide a comparative analysis of the perceptual representation of the vowel space of AE as entertained by nonnative PA learners (group A) and by native speakers of AE (Group B) through an identification task of synthesized vowels. In Chapter 5 (Study 3), I test the 11 AE monophthongs as produced by L1 PA learners via a production task and statistically compare their results with production results available for native AE speakers. In Chapter 6, my findings are critically compared and evaluated internally across the three studies and externally (where applicable) with the results of other L1 Arabic EFL similar studies in relation to the reviewed body of literature and the questions proposed for each study.

Each of the three experimental studies represents a standalone study with sections covering more specific questions, hypotheses, related literature review, methodology, results, and discussion sections. In the conclusion chapter (Chapter 7), however, I will recapitulate the general as well as the more specific research questions and hypotheses, then summarize my main findings and try to answer the general questions asked. Moreover, in the Conclusion Chapter, I will suggest new research questions that may have been prompted by my research and identify questions that could not be answered yet, either due to lack of time or methodological shortcomings.

CHAPTER TWO

LITERATURE REVIEW

2.1. Contrastive analysis of languages

This section of the chapter presents a general overview of Modern Standard Arabic (MSA), Palestinian Arabic (PA), and American English (AE). For each language (or variety), a discussion will be provided on the vowels, consonants, and consonant clusters, while additional details about suprasegmental (syllabification and stress) features will follow. Finally, there will be a comparative analysis between PA and MSA at the phonemic level to show how deviant the PA is from MSA. Additionally, a contrastive analysis will then be provided between AE and PA based on the similarities, differences, and alternations across these languages. The comparison will base my predictions and hypotheses on any type of pronunciation or perception errors that PA EFL learners might commit based on the degree of (dis)similarity at the phoneme level. The section is overlaid with an overview of the main differences at the phonological level and predictions of the major phonotactic constraints and learning problems.

2.1.1.Modern Standard Arabic (MSA)

It is said that each language has its own particular phoneme inventory that contains the contrastive sounds of that language. Starting with MSA, many scholars highlight the idea that MSA is a language with a small vowel inventory and a rich consonant inventory (e.g., Cohn, 2001: 182; Flemming, 2001: 12; Holes, 2004: 1; Mustafawi, 2018: 12; Watson, 2002: 21). Additionally, for a strictly phonemic orthography or a phonologically transparent orthography, the phonemes in a language should correspond in a one-to-one relationship with their orthographic symbols, i.e., (one sound = one symbol). This is mainly true for Arabic, with some minor exceptions. In MSA, there are 28 consonants with their corresponding 28 orthographic symbols, three long vowels with orthographic symbols, and three short vowels that can be expressed by diacritics but are not generally used in everyday writings since native Arabic speakers can infer the correct pronunciation from the context. In addition, Arabic has two diphthongs represented by two adjacent vowels. On the other hand, the orthography of English is not transparent in regard to grapheme-phoneme correspondences. In contrast, it is considered to have a complex and divergent (nontransparent or opaque) spelling system reflected in many discrepancies in the written form, e.g., silent letters, digraphs, split digraphs, trigraphs, tetragraphs (e.g., *b<u>ough</u>t* /b<u>ɔː</u>t/), or one letter representing more than one sound. EFL learners' perception and production of the sounds in the target language may be compromised if their L2 acquisition is unduly guided by their assumption that the spelling-to-sound correspondence is as regular as in their native language. To avoid such negative transfer, explicit instruction is called for if EFL learners' L1 has a transparent spelling system, such as Arabic. However, this is just one part of the bigger problem that is rooted deeply at the phoneme level. Therefore, the following subsections will detail the characteristics of two varieties of Arabic (i.e., MSA and PA) and one variety of English (i.e., AE) at the segmental and suprasegmental levels, starting with those of MSA.

2.1.1.1. Vowels

MSA has a three-point vowel system, which is one of the simplest and most straightforward types. In the orthography of many languages, vowels that contrast by duration are given the same symbol, but the long ones are indicated by diacritics superimposed on the orthographic symbols. In contrast, Arabic does the opposite. Basically, the three short vowels of MSA, /i, u, a/, are indicated in writing using diacritics only (above or below preceding consonants). These short vowels are doubled in number and contrast phonemically by duration to provide three long counterparts /i:, u:, a:/. As a result, there are six vowels /i, i:, u, u:, a, a:/ and two rising diphthongs /ai, au/ (Al-Ani, 1970: 22; Mitchell, 1993: 138; Watson 2002: 22–23; Newman & Verhoeven, 2002: 79; McCarus, 2011: 527-528; Embarki, 2013: 7) that contrast along three parameters: (1) the height of the body of the tongue, i.e., high vs. low, (2) the front-back positioning of the tongue, i.e., front vs. back., and (3) duration, i.e., short vs. long.

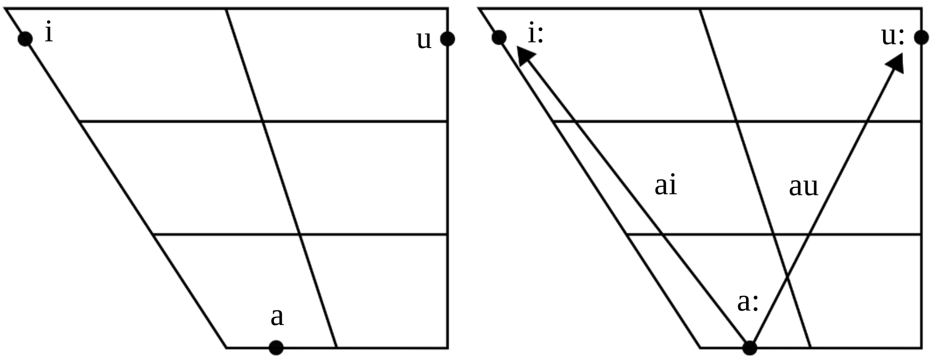

Figure 2.1. MSA short, long, and diphthong vowels. After (Thelwall & Sa'adeddin, l999: 52).

Figure 2.1 showcases MSA vowels and their distribution over the IPA vowel chart. Accordingly, the MSA vowels can be described as /i:/ high front long (/ʕi:d/ 'holiday'), /a:/ low central long (e.g., /ba:b/ 'door'), /u:/ high back long (e.g., /nu:m/ 'sleep' (noun)), /i/ high front short (e.g., /ʕiqd/ 'a decade'), /u/ high back short (e.g., /ʕuqd/ 'a necklace'), /a/ low central short (e.g., /ʕagd/ 'contract'), /ai/ low central short to high front long (e.g., /bai:t/ 'house'), and /au/ low central short to high back long (/khau:f/ 'fear').

According to Figure 2.1, the vowels /i/ and /i:/ are high front vowels, /u/ and /u:/ are high back, while /a/ and /a:/ are low central vowels, and each set differs only in duration, whereby the long counterparts are marked with /:/ after the base symbol in the transcription. Diphthongs count as long vowels. Lip rounding is not contrastive in Arabic; high back vowels are produced with (slight) lip rounding.

An argument concerning vowels in Arabic phonology is whether the MSA long-short vowel contrast crucially involves quality (correlated with formant structure) as phonetic correlates. In terms of vowel quality, some researchers argue that short vowels do not yield the same exact quality as their longer counterparts. Al-Ani (1970) reported that there are quality differences between short-long counterparts and not only durational differences: "It appears that with /a/ and /aa/ there is not only a quantity difference but a substantial quality difference as well." (p. 24).[3] Watson (2002: 22) found that the articulation of both /iː/ and /uː/ Arabic long vowels is higher than that of their short cognates and that /aː/ has fronter articulation than its short counterpart in all Arabic dialects. The double marking of the contrast (in duration and vowel quality) would be an argument to consider the contrast as a tense-lax opposition rather than just a length opposition.

2.1.1.2. Consonants

I mentioned earlier that Arabic has a rich consonant inventory that includes 28 consonantal phonemes (29 if the glottal stop /ʔ/ is given phonemic status, Al-Ani, 1970: 29; Holes, 2004: 58). The spelling and pronunciation variants of the Arabic glottal stop [hamza] are among the very few exceptions to the overall transparency of grapheme-phoneme consonants in Arabic. Basically, there are two types of pronounced glottal stop (or /hamza/): weak and strong. The strong (or severing) glottal stop is called [hamza al-qatˤʕ], which is phonemic and is pronounced under all circumstances, whether word-initially, -medially, or -finally. This sound is similar to the catch in the voice between the syllables of "oh-oh". The weak (or elidable) glottal stop is [hamzat al-wasˤl], which functions as a phonetic carrier that "helps pronunciation of consonant clusters and only occurs at the beginning of a word" (Ryding, 2005: 16), is often deleted in notation.[4]

This rich consonantal inventory places Arabic above the average of the world's languages, which is (22 ± 3) consonants according to the World Atlas of Language Structures Online (WALS) (Maddieson, 2013a). What distinguishes Arabic even more, as Table 2.1 showcases,

[3] The notation /aa/ is used in the reference to represent two morae, the same as /a:/ for Arabic in this thesis.
[4] For further discussion on the glottal stop variants in MSA, the reader is referred to Ryding (2005: 14–16) and Testen (1998: 135–138).

is the presence of uncommon phonemic consonants that are not available in most other languages. According to the WALS, up to ten consonants, such as pharyngealized fricatives and stops (i.e., /ħ, ʕ/ and /tˤ, dˤ/), exist in Arabic but not in the most widespread languages, such as English or French. As detailed in Table 2.1, in the case of a voiced/voiceless contrast, consonants to the top of a cell are voiced, and those to the bottom are voiceless. Rows without voiced/voiceless contrast contain voiced consonants only.

Table 2.1. MSA consonantal system. Based on (Embarki, 2013: 39; Mustafawi, 2018: 12; Ryding, 2014: 15; Watson, 2002: 20).[5]

MSA	Labial	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	b		d dˤ t tˤ			k	q		ʔ
Nasal	m		n						
Fricative	f	ð ðˤ θ	z s sˤ	ʃ		ɣ ~ x ~	ʁ χ	ʕ ħ	h
Affricate				ʤ					
Tap/Trill			r						
Lateral			l						
Glide	w				j				

Key: the diacritic "ˤ" denotes the emphatic consonant. Phonemes in the upper half of a row are voiced.

There is one more main difference that distinguishes Arabic from English, that is consonants with secondary articulations (the velarized, pharyngealized, or "guttural" consonants); the velar and pharyngeal strictures produce distinctive consonants in Arabic, such as the voiceless uvular stop /q/, the phonemic glottal stop /ʔ/, or the emphatic pair of dental voiced and alveolar voiceless fricatives /ðˤ/ and /sˤ/, respectively. These sounds can be problematic if the case were inverted and Arabic is the L2 while AE L1 learners need to learn how to pronounce these new sounds that involve secondary articulations that AE lacks. The emphatics are unusual sounds in the languages of the world. They are marked in the sense of Eckman's Markedness Differential Hypothesis (1977). The marked version will therefore not be used by Arabic EFL learners because English only has non-emphatic consonants (i.e., the unmarked versions). Therefore, there will be no interaction in the mind of the EFL learner. The emphatics will only be a problem for learners of Arabic (if the learner's L1 has no emphatics). There are other possible contrasts between the two languages, but their discussion is deferred to the following contrastive analysis section (Subsection 2.3.4.2).

[5] For a full account on MSA consonant stricture places, the reader is referred to Ryding (2014: 15-16) or Mustafawi (2018: 14-16) or more exhaustively to Watson (2002: 14-19) or Al-Ani (1970, Ch 1).

2.1.1.3. Syllable structures and consonant clusters

The MSA formalization of syllable structures in a metrical system can be embodied in this schema: CV(:)(C)(C) (Mitchell, 1993; Yavaş, 2011: 197).[6] Any syllable in MSA starts with one consonant C, which is obligatorily followed by a short or long vowel V(:) and generates an open syllable in this case. Both short and long vowels can be followed by maximally two Cs in coda to create a closed syllable. MSA has one light syllable type (CV), two heavy syllable structures (CV: and CVC), and three superheavy syllables (CV:C, CVCC, and CV:CC). MSA syllables are constrained further by rules concerning the pronunciation form, which is divided into two types: full-form pronunciation (continuous speech) and pause-form pronunciation, which omits the final short vowel in the prepausal position. In full-form pronunciation, only light and heavy syllables are allowed, e.g., words such as /ʕa:li:/ 'high' or, /madrasah/ 'a school' are pronounced /ʕa:.li:/ and /mad.ra.sah/, respectively. Additionally, one superheavy syllable is possible in full-form pronunciation, viz. CV:C, but in restricted circumstances within an MSA word. Such situations most often resulted from a morphophonological process to the word root according to Ryding (2014: 35). For example, /ma:d-da-tun saʕ.ba.tun/ 'a difficult course' or /ʔa.da:.tun ħa:d.dah/ 'a sharp tool'. In pause-form pronunciation, all of the six syllable forms are allowed. The superheavy syllable structures CVCC and CV:CC appear only in pause-form by omitting final short vowels (Ryding, 2005: 39; 2014: 35) when the speaker stops for a pause, e.g., /ʕitˤr/ 'perfume' or /bint/ 'girl', and in the case of CV:C, thereby creating a word-final geminate: e.g., /tˤa:ll/ 'prolonged', /ʒa:dd/ 'serious', and /ʕa:dd/ 'returned'. An example to compare full-form and pause-form speech would be, for instance, [sa:ʔiqu ʔalħa:felah] 'bus driver, masc.' can be pronounced as /sa:.ʔi.qul.ħa:.fe.lah/ in full-form, but if the speaker stopped after the first word, the correct pronunciation becomes: /sa.ʔiq-ʔal.ħa.fe.lah/. It is also clear from this example that syllable boundaries in connected MSA speech are not necessarily coterminous within word boundaries.

Based on the schema above, two rules of interest to this thesis can be quickly inferred. First, MSA does not allow an onset-free syllable. Therefore, whenever a word or a syllable starts with a vowel (i.e., zero-onset), phonetically, it is pronounced with an epenthetic glottal stop /ʔ/. Second, Consonant Clusters (CCs) are not allowed word-initially (i.e., more than one C in word-onset is not tolerable in MSA). By extension to superheavy syllables, CCs within syllables are also prohibited, except in pause-form and at the end of a word.

[6] By convention, constituents in parentheses are optional. They can be either included or omitted.

In colloquial Arabic, these scenarios vary according to the L1 dialect. Cairene Arabic adds an epenthetic vowel word-initially, e.g., MSA /qul.tu.lah/ ‘I told him’, becomes /ʔul.ti.lu/ in Cairene Arabic (with /q/ turning into /ʔ/), while in the Iraqi dialect, it is pronounced as /gə.lit.lu/ and in Sudanese /gə.li:t.lu/. This pattern is recurrent and difficult to suppress in Arabic-accented English in the realization of words beginning with two or three consonant clusters.

2.1.1.4. Stress

One of the structural features that ensure that speech is not monotonous or level is stress. Stress as a phonetic phenomenon is discussed in more detail in Subsection 2.1.3.4. A discussion concerning the abstract status of stress in Arabic is presented here.

Many historical studies have dealt with Classical Arabic (CA) grammar, but in regard to stress, there are no reliable resources in these studies on what kind of stress it followed. Mion (2011) clarified this situation with two remarks. First, it is argued that Arabic grammarians focused more on written forms and the correct pronunciation of consonants. Second, it is also argued that more phonological value is specified for vowel quantity rather than to stress as a contrastive feature in CA. This means that stress is not used to indicate differences in lexical meaning or different grammatical classes. Additionally, Watson (2011) justified the poor body of literature by old grammarians on Arabic stress by suggesting that stress assignment was largely predictable. Indeed, contemporary studies on Arabic stress (Astruc, 2013; Denham & Lobeck, 2013, among others) have shown that MSA (and its colloquial varieties) is a weight-sensitive stress-timed language (like English) with fixed and regular word stress (unlike English) (Ryding, 2005 on MSA; Ghazali et al., 2007 on CAVs; Al-Mallah, 1983; Davis, 1995; Hall, 2017; Monahan, 2002 on PA). The absence of stress in old studies may be due to the fact that stress was not contrastive in MSA (Mustafawi, 2018). However, even with a lack of evidence on CA having exhibited stress, many researchers (e.g., Angoujard, 1990; Fischer 1997; Holes 2004: 61; McLoughlin, 2009:5) agree on a variety of acceptable stress patterns for MSA that adhere to the following rules. In the full-form pronunciation, stress falls on the second or third syllable from the end of the word, i.e., stress the penultimate syllable of the word if it is heavy, otherwise stress the antepenultimate syllable). In the pause-form pronunciation, the stress rules remain the same, but one rule is added for the superheavy syllables, i.e., stress the final syllable of words ending in CVCC or CV:C.[7] In summary, as reported by McCarthy and

[7] For a full account of MSA stress patterns, see Mion (2011); Ryding (2005, 2014: 33-36) and Mustafawi (2018: 20-22)

Prince (1990: 252) and confirmed in Ryding (2005, 2014), MSA stress falls on the final syllable if superheavy, on the penult syllable if heavy, and on the antepenult syllable otherwise.

2.1.2.Palestinian Arabic

It is claimed that there are no native speakers of MSA, and speakers of Arabic currently speak their countries' colloquial varieties of the language (Badawi, 1996; Kaye, 2001). Generally, CAVs differ from MSA in many aspects, such as having a reduced and restructured consonant system and more developed vowel inventories (on the surface level). Moreover, CAVs have different syllable structures than MSA, which accordingly affect the overall formation of words and lexical stress (Watson, 2002). In the following section, I will detail the specifications of PA at the segmental and suprasegmental levels.

2.1.2.1. Vowels

Earlier, it was discussed that there are three phonemic qualities in the MSA vowel inventory, which contrast in duration to yield six short and long vowels /a, i, u, aː, iː, uː/ and two diphthongs (/ai/, /au/). However, the Arabic variety spoken in the Levantine, which includes Palestinian, Jordanian, Syrian, and Lebanese varieties, has been analyzed with many more vowels. For example, the Syrian Arabic variety is reported to have eleven different phonemic vowels /iː, eː, aː, oː, uː, i, e, a, o, u, ə/ (Almbark & Hellmuth, 2015). More particularly for this research, PA was found to have five long vowels /iː eː aː oː uː/ and at least three short vowels /i, a, u/ (Hall, 2017). Other studies in this context (Blanc, 1953; Palva, 1965; Tsukada, 2009; Shahin, 2011; Amir et al., 2014) considered PA to have up to ten or eleven vowels, as shown in Table 2.2 below.

Table 2.2. Palestinian Arabic vowels. Based on Shahin (2011: 529).

	Front		Back
High	i iː		u uː
Mid	e eː		o oː
Low		a aː	

As seen in Table 2.2, PA vowels differ in height and backness as well as duration, just as in the MSA vowel system. However, some vowels in the PA vowel system are reportedly not shared with MSA, i.e., at least two long vowels /eː/ and /oː/ (Hall, 2017: 4) or two additional short vowels /e/ and /o/ (Amir et al., 2014: 1; Palva, 1965: 171; Shahin, 2011: 529). However, concerning the allophonic variation in the PA variety, Palva (1965: 9) clearly stated that "There are only three phonemic vowel qualities, *a, i,* and *u*. The number of their phonetic allophones is indefinite and it seems to me useless to mark more than these eleven allophones" (p. 9).

Cowell (1964) mentioned that Lebanese and Palestinian Arabic do not contrast between /o/ and /u/ or between /a/ and /ə/. As mentioned, these additional vowels do not have a contrastive role within PA in general. They are nonphonemic and do not change the meaning of a word and only represent cases of free variation, e.g., [bint] or [bɛnt] ‘a girl’. I hinted earlier that the MSA vowel system has many allophones for the original six vowels. Concerning this, Al-Ani (1970: 23–24) and Canepari (2007: 317) listed up to 17 different allophonic realizations of the six MSA vowels. In the same vein concerning PA, the point vowels /i, a, u/ and their long counterparts /iː, aː, uː/ are reported to have several allophones in PA. Most usually, PA [ʌ, ə, e] are allophones of MSA /i/; PA [o] is an allophone of MSA /u/; PA /eː, oː/ of MSA /ai, au/ (Palva, 1965: 8-9). Therefore, in this study, as in Saadeh (2011), PA is considered to have the six basic vowels – as in MSA.

2.1.2.2. Consonants

The PA consonant inventory can have up to 30 consonants. Most of them are shared with MSA, and some of them have an allophonic status that varies according to local PA dialects (i.e., urban, rural, Bedouin).

Table 2.3 shows the PA consonant inventory while accounting for the differences at the regional dialect level. Additionally, the major allophonic differences in PA consonants among the three local dialects are detailed in Table 2.4 in comparison to MSA consonant counterparts.

Table 2.3. PA Consonantal System. Based on (Cotter, 2020; Shahin, 2011; Palva, 1965).

PA	Labial	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	b		d dˤ t tˤ			+g k	*q		ʔ
Nasal	m		n						
Fricative	f	*ð *ðˤ *θ	z s sˤ	ʒ ʃ		ɣ ~ x ~	ʁ χ	ʕ ħ	h
Affricate				**ʤ** tʃ					
Tap/Trill			r						
Lateral			l lˤ						
Glide	w				j				

Key: * Consonant does not occur in urban dialect. + Consonant occurs only in the Bedouin dialect. **Boldface**: consonant occurs only in rural and Bedouin.[8]

As shown in Tables 2.3-4, Arabic consonants in PA have changed from MSA in number and pronunciation, similar to all other CAVs.

[8] For a full account on PA consonants, see Shahin (2011: 528).

Table 2.4. PA Consonant's cognates and dialects displacement. After (Cotter, 2020; Shahin, 2011).

Language	Consonants						
MSA	k	q	θ	tˤ	ð	ðˤ	h
Urban PA	k	ʔ	t	t	d	{dˤ, ðˤ}	{w, u:}
Rural PA	{tʃ, k}	k	θ	tˤ	ð	ðˤ	{w, u:}

PA dialects have introduced new consonants relative to MSA, e.g., /tʃ/ in rural PA and /ɡ/ in Bedouin PA; both act as allophones of MSA consonant /q/, which in urban PA became /ʔ/. To illustrate this case, consider the word /qalb/ 'heart' in MSA. The colloquial versions thereof in PA are /ʔalb/ in urban PA, /kalb/ in rural PA, and /ɡalb/ in Bedouin PA. Most of the consonant changes in urban PA have been in the pharyngealized, such as /q, tˤ, ðˤ/, which have shifted to their nearest counterparts /ʔ, t, z/, respectively. Finally, the /h/ sound is replaced by /w/ or /u:/ word finally.

2.1.2.3. Syllable structure and consonant clusters

There is a tendency in CAVs and PA in particular (Abu-Salim, 1982, 1986; Hall, 2017; Palva, 1965: 16; Shachmon, 2017) toward either preserving, omitting, or reducing short phonemic vowels (especially in open unstressed syllables in case of reduction) and forming new nonphonemic vowels that affect the syllable structure and yield new syllables with consonant clusters. Therefore, the syllables in the PA variety can be captured by the following formula (C)(C)V({:, C})(C) (Shahin, 2011).[9] Accordingly, PA syllables are either (1) open syllables: CV, CCV, e.g., /bti.ʒi:/ 'you come, 2nd masc.', CV:, CCV:, e.g., /bta:.kul/ 'you eat, 2nd masc.' or /nʒa:.sˤa/ 'a pear' or (2) closed syllables: CVC, CCVC, e.g., /btel.ʕab/ 'you play, 2nd masc.' or /bsur.ʕa/ 'quickly', CVCC, CCVCC, e.g., /nzilt/ 'I went downstairs' or /lʕibt/ 'I played', VC, VCC, V:C (word-initially) are all preceded by an epenthetic glottal stop, e.g., the MSA /yadd/ 'a hand' is /ʔi:d/ in PA, and the CV:C, CCV:C, e.g., /bla:d/ 'countries', or /ʕru:q/ 'veins'.[10]

As seen from the schema above, there are many more syllable structures possible in PA than in MSA. Some of these include word-initial consonant clusters only, and other syllables happen everywhere else in all possible positions, i.e., word-initially, -medially, and -finally. More specifically, CC clusters in heavy syllables are allowed word-initially as in /mʕal.lem/ 'teacher, sing. masc.', /msar.rab/ 'leaked', or word-medially /ʃuft.hum/ 'I saw them'. Word-final CC occurs only phrase-internally or when C1 is /r/ (or any other approximant), e.g. /ħa.dart.id.dars/ 'I attended the lecture'. Word-initially, CC may vary freely with VC epenthetic syllables, e.g.,

[9] By convention, constituents in curly brackets are disjunct options: one and only one can be chosen.
[10] Only syllables not possible in MSA are exemplified, and underlined in polysyllabic words.

MSA /ħi.sˤaːn/ 'a horse' is pronounced either as /ʔiħ.sˤaːn/ or /ħsˤaːn/, interchangeably in PA (Palva, 1965; Shahin, 2011; Tayeh et al., 2012).

Studies on CAVs propose that the typology of variation in syllabification differs from one variety to another. Instead of the two-way crosslinguistic typology of rhythm, many studies (e.g., Palva, 1965) suggest a continuity of rhythmic variations across languages emerging from independent variables editing the syllabification, such as (extreme) vowel reduction. Although CAVs are stress-timed, studies have shown that they vary between more to less stress-timed between the east (Levant) and west (Morocco) (Ghazali et al., 2002; 2007; Kenstowicz, 1986). This surface variation phenomenon coincides with the allowed syllable types and prosodic structures found in different CAVs (Embarki, 2013; Hamdi et al., 2005; Hellmuth, 2013). Accordingly, the different consonant clusters in PA can be found in some neighboring varieties by extension, which may not necessarily occur in other CAVs. Therefore, studies on other Arabic varieties can reveal other possible CC features.

2.1.2.4. Stress

CAVs have received more attention in verifying clear stress characteristics and patterns than MSA. Studies have highlighted the issue of whether long and short vowels in different CAVs have a different gradation of acoustic correlates, i.e., fundamental frequency (F0), duration, and intensity. In relation to F0, it is found that if stressed syllables, in a certain language, are marked with an increment of the F0 values in general when stress is called 'musical', while if the stressed syllable is marked with an increment of the duration of its vowel and/or intensity, then the stress is called either 'temporal' or 'dynamic' (Van Heuven, 2002). According to Lehiste (1970), languages are divided into two groups according to their stress systems: languages with fixed stress in words, such as Hungarian, French, or Finnish, and languages with non-fixed (variable) stress that can change its place in a word and accordingly change the meaning, as in English, German, and Spanish. English stress is discussed in section 2.3.2.4, but concerning Arabic varieties, Mion (2011) reported that the majority of CAVs have a musical stress type, i.e., the primary lexical stress in these varieties is manifested by an increment of F0 values on the nuclear vowel. In contrast to other languages (e.g., Italian), lexical stress in Arabic does not affect the duration of the vowel but increases the tonal height (and intensity) of the vowel.

At the phonological level, the available data are abundant on different colloquial Arabic stress patterns (e.g., Hellmuth, 2013; Holes, 2004; Mustafawi, 2018; Watson, 2002 and references therein). Such studies confirm that modern dialects of Arabic abide by the assumed rules of MSA stress and that in all Arabic dialects, stress is weight-sensitive (quantity sensitive).

As for Palestinian Arabic in full-form pronunciation, stress patterns are as follows: never stress the final syllable, stress the prefinal superheavy syllable (CV:C or CVCC), and stress the antepenultimate in all other cases (Hellmuth, 2013). Additionally, the MSA stress rules in pause-form apply to PA pronunciation and remain the same (i.e., a final superheavy syllable receives the main stress). However, Kenstowicz and Kisseberth (1979: 229–231) proposed different stress rules for PA. Generally, if the final syllable contains a long vowel or ends in a consonant cluster, it receives the stress; otherwise, stress is prefinal if heavy and antepenult in all other cases. There are some surface exceptions to these PA stress rules. Sometimes stress is antepenultimate even if there is a penultimate heavy syllable. Another exception is to stress the final even if it does not have a long vowel or a consonant cluster.

To provide a few examples of PA stress in comparison with other non-Levantine Arabic varieties (e.g., Cairene Arabic), it is noted that in Levantine Arabic, if the final syllable is superheavy (CV:C, CVCC), it receives the primary stress in pause-form, as in PA: /jaˈwa:b/ 'answer', /biˈnutˤtˤ/ 'he jumps', which is the same as in Cairene Arabic as in /fiˈlu:s/ 'money', /hiˈdu:m/ 'clothes'. In case of no final superheavy syllable, stress is allocated to heavy non-final (CV: or CVC), as in PA: /musˈtaʃ.fa] 'a hospital', /muˈna:.fis/ 'a competitor', which is also true for Cairene Arabic: /daˈras.tu/ 'you (pl) studied', or /ʃirˈbu:.ha/ 'they drank it'. Meanwhile, in cases with no final superheavy or a heavy penult, stress differs according to the variety. Such that in Cairene [madˈra.sa] 'a school', contrasts with PA [ˈmad.ra.si]. Finally, words with light syllables only receive the stress on the initial syllable, e.g., [ˈka.tab] 'he wrote' or [ˈʃa:f.hum] 'he saw them' (Abdulkarim, 1980; Adra, 1999; Alghazo, 1987; Mion, 2011).

It should be clear by now how CAVs, and especially PA, stress patterns and rules diverge partially from those of MSA and clearly differ from those of English (see section 2.1.3.4). The failure to process these differences will lead to a foreign accent, as will be further discussed in the contrastive analysis section at the end of this chapter (section 2.1.4.4).

2.1.3. American English

This section will start by describing AE vowels, consonants, syllable structures, and stress patterns while providing examples for ease of comprehension. Then, it will elaborate on the acoustical properties of vowels to practically provide measurable results.

2.1.3.1. Vowels

In comparison with other languages, English has a large inventory of vowels. According to the WALS, the English vowel inventory is placed well above the vast majority of the world's natural languages that contain 5 or 6 vowel phonemes (Maddieson, 1984: 127; 2013b).[11] These differences, when put in a language acquisition context, pose serious challenges for learners whose native languages lack one or more of the English vowels since their abilities to perceive and produce them intelligibly are affected. This prompts questions such as how EFL learners will replace the missing L2 vowels with vowels of their L1 inventory, what compensating strategies will they develop, and how these strategies interfere with the segmental intelligibility of the produced utterances (Maddieson, 2013a; Schwartz et al., 1997).

The AE inventory has 15 vowel sounds in total (Labov et al., 2006: 14-15; Ladefoged, 1999: 42; Wells, 1982: 120, 472). These are /ɪ, i, ɛ, æ, ʌ, ʊ, ɑ, u, e, o, ɔ, aɪ, ɔɪ, aʊ, ə/. These vowels are listed for the sole purpose of manifesting the complexity of AE in comparison to Arabic, but as mentioned earlier, the present dissertation explicitly excludes vowel reduction and schwa from the present dissertation as these phenomena do not occur in stressed monosyllables (see 2.1.3.4). Therefore, the perception and production studies in this dissertation will only tackle the 11 AE monophthongs, i.e., /i, ɪ, e, ɛ, æ, ʌ, u, ʊ, o, ɔ, ɑ/ (excluding schwa /ə/ and schwar /ɚ/).[12] I will start by describing two main characteristics of the vocal tract involved in vowel production. First, vowels are described according to backness (front, middle, or back) and height (high, mid, or low). Then, I elaborate on the tense-lax categorization of AE in relation to their centralized vs extreme spectral distribution. Figure 2.2 shows how the AE monophthongs (A.) and diphthongs (B.) are distributed over the vowel chart. The inner polygon in panel A connects the lax vowels /ɪ, ɛ, ʌ, ʊ/; the other monophthongs, i.e., /i, e, æ, u, o, ɔ, ɑ/ are phonetically tense.

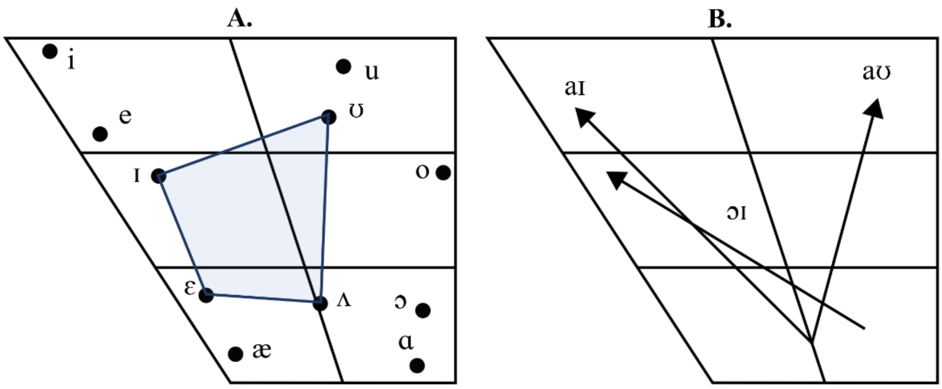

Figure 2.2. AE monophthongs (A.) and diphthongs (B.). After Ladefoged (1999: 41).

In addition to the previous characteristics of vowels, the shape of the lips (lip-rounding) while producing the vowel is another parameter directly involved in describing vowels, i.e.,

[11] For the relevant section on American English vowels in the World Atlas of Linguistic Structures (WALS), see https://wals.info/feature/2A#2/19.3/152.8

[12] For further discussion on r-colored and reduced vowels, see Wells (1982, vol. III, Section 6.1.5: 479–485)

rounded monophthongs /u, ʊ, o, ɔ/, and unrounded monophthongs (the rest). In AE (and other varieties of English), lip rounding is redundant: Only non-low back vowels have rounded lips; all other vowels are unrounded. Table 2.5 below provides a detailed description of the AE monophthongs according to the three basic parameters of vowel quality while providing other symbols used interchangeably to replace certain IPA symbols and exemplifying words for each vowel (based on Ladefoged, 1996; Maddieson, 1984; Yavaş, 2011; Wells, 1982).

Table 2.5. AE monophthongs in detail.

Symbol	Alternative symbols	Description	Example
i	/iː/, /ij/, /iy/	high front unrounded	peat, near
ɪ		near-high near-front unrounded	pit, kit
e	/eɪ /, /ej/, /ey/	high-mid front unrounded	lay, face
ɛ		low-mid front unrounded	pet, dress
æ		near-low front unrounded	pat, trap
ʌ	/ə/ (*unstressed syllables*)	mid-central	but, strut
u	/uw/, /uː/	high back rounded	shoe, goose
ʊ		near-high near-back rounded	put, foot
o	/oʊ/, /ow/	high-mid back rounded	low, goat
ɔ	/ɔː/	low-mid back rounded	law, thought
ɑ	/aː/	low back unrounded	caught, lot

Another important vocalic feature that contrasts AE vowels is the tense/lax parameter. Some languages, such as Dutch, Frisian, and English, have a phonological opposition of tense (long) vs lax (short) vowels (Chomsky & Halle, 1968: 324; Labov et al., 2006: 14; Zsiga, 2013: 60) as a phonemic parameter for the vowel inventory. In this dissertation, the tense-lax parameter is used in the sense that tense vowels are described to have a more extreme articulatory position (more peripheral tongue positioning), more tension of the tongue and lip muscles and articulatory effort than their 'lax' counterparts, and they are longer in duration (Fromkin et al., 2019: 202; Labov et al., 2006: 16; Stevens, 2000). The division of the 11 AE pure vowels into 4 short and 7 long vowels is undisputed (e.g., Celce-Murcia et al. 2010; House, 1961; Lehman & Heffner, 1940; Peterson & Barney, 1952; Wang & Van Heuven, 2006). Most authors observe that the short vowels assume more centralized articulations in the vowel space, while the long vowels are articulated on the outer perimeter of the space. On the assumption that it takes more muscle activity to move the articulators to extreme/peripheral positions, some authors call the peripheral vowels 'tense' and their more centralized counterparts 'lax' (e.g., Celce-Murcia et al., 2010: 114-115; Koffi, 2021: 9; Strange et al., 2004; Wang & Van Heuven, 2006; Yavaş, 2011: 79). There is even some physiological backing of greater muscle tension for the peripheral vowels in American English (Raphael & Bell-Berti, 1975).[13]

[13] The phonetic tense-lax distinction does not completely parallel with the distributional observation that tense vowels can occur in open syllables in English, while lax vowels cannot, and that tense vowels cannot be followed

Another major characteristic that distinguishes between different American English dialects is the low-back merger, which is also known as the *cot-caught* merger. The vowel /ɔ/ is pronounced (and perceived) as /ɑ/ in the dialect where this merger occurs. Therefore, words such as *cot-caught* or *stock-stalk* are pronounced the same. Wells (1982) mentioned that this distinction is a well-known diagnostic tool to distinguish different dialects of AE, as it prominently spreads, and is more persistent, in the northern speech area of the US than in the midland and southern areas. Furthermore, it is more prominent in the Western USA dialects than in the Eastern USA dialects (except the northeastern dialect), e.g., the US Southern accent (Texan English) or the Miami accent (Labov et al., 2006). The relevance of this merger to the present research is that the majority of the control speakers and listeners used in the present dissertation have lost the low-back contrast.

2.1.3.2. Consonants

The AE consonant inventory with 24 consonants is slightly above the mode (22 consonants) among the world's natural languages (WALS database, Maddieson, 2013a). Primarily, American-English consonants /p, b, t, d, k, g, ʧ, ʤ, f, v, θ, ð, s, z, ʃ, ʒ, m, n, ŋ, l, r, w, j, h/ contrast according to their manner of articulation, place of articulation, and voicing. AE consonants employ eight different places of articulation. These are: 1. Bilabial / p, b, m, w/, 2. Labiodental /f, v/, 3. Interdental/ θ, ð /, 4. Alveolar /t, d, n, s, z, ɹ, l/, 5. Post-alveolar /ʃ, ʒ, ʧ, ʤ/, 6. Palatal /j/, 7. Velar /k, g, ŋ, /, 8. Glottal /h/. Also, AE is commonly described with six manners of articulation. These are: 1- Stops (or plosive) /p, b, t, d, k, g/, 2. Fricatives /f, v, θ, ð, s, z, ʃ, ʒ, h/, 3. Affricates /tʃ, ʤ/, 4. Nasals /m, n, ŋ /, 5. Liquids /r, l/ 6. Glides /w, j/. The voicing feature in AE can split the consonants into two types; voiced /b, d, g, ʤ, v, ð, z, ʒ, m, n, ŋ, l, r, j, w/ and the voiceless consonants /p, t, k, ʧ, f, θ, s, ʃ, h/ are produced with the vocal folds spread as in the position of normal breathing.

Table 2.6. AE consonants. After Ladefoged (1999: 41).

AE	Bilabial	Labio-dental	Interdental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	b p			d t			g k	ʔ**
Fricative		v f	ð θ	z s	ʒ ʃ			h

by a coda /ŋ/. The parallel is perfect in British English. In AE, the situation is somewhat diffuse since /æ/ and /ɑ/ are phonetically tense (and long) and yet occur in closed syllables that may even end in /ŋ/, as in *fat*, *fang*, and *lot*, *long*.

Affricative			dʒ tʃ		
Nasal	m	n			ŋ
Liquid		ɹ l			
Glide	w*			j	w*

Key: * /w/ is labial-velar with two stricture places of open approximation: labial and velar. ** /ʔ/ English has the segment [ʔ] but only as an allophone of /t/- and /d/-glottalization in words such as *wait* [waiʔ], *button* [bʌʔn] and *hidden* [hɪʔn].

For comparison, Table 2.6 provides a detailed categorization of AE according to their place, manner, and laryngeal features (voicing). In each row with consonant pairs, the consonants at the bottom of a cell are voiceless, and those at the top are voiced. Cells without voiced/voiceless contrast contain voiced consonants only (non-obstruents).

English sounds can further be categorized as obstruents (a.k.a. true consonants) vs. sonorants. English obstruents, which are produced with an obstruction in the vocal tract that prevents spontaneous voicing, include stops, fricatives and affricates. Meanwhile, English sonorants, which have a less severe obstruction so that the airstream is almost unimpeded and still allows spontaneous vocal fold vibration (as in vowels), include nasals, liquids, and glides. Obstruents are naturally voiceless, but marked obstruents can be produced with voicing if the speaker expends extra effort to allow the vocal folds to continue to vibrate (e.g., expand the supraglottal cavities). Accordingly, the obstruents come in voiced-voiceless pairs that share all parameters but differ in voicing, e.g., /f-v/, /b-p/, and /t-d/, etc. Sonorants (/m, n, ŋ, l, r, w, j) are always voiced, and have no contrastive voiceless counterparts. The voiceless sound /h/ has no voiced counterpart (Ladefoged & Disner, 2012). This categorization will be handy in the discussion on AE consonant clusters in section 2.1.3.3 below.

Most AE consonants occur in all positions of the word: at the beginning, in the middle, or at the end. However, there are some limitations. Some consonants are not possible in certain positions, e.g., the /ŋ/ sound is not possible word-initially but possible in the other positions, and the /h/ sound cannot occur word-finally (nor can the semivowels /j, w/).

2.1.3.3. Syllable structure and consonant clusters

The syllable structure of English is complex compared to Arabic. The structure of English syllables consists of onset (any segments preceding the vowel) and rhyme (the remaining part of the syllable). The onset in English can contain either one, two, three prevocalic consonant(s), or none. The syllable rhyme consists of two parts: a nucleus and a coda. The nucleus is an obligatory member of the syllable and typically contains a vowel. However, it is possible for a syllabic consonant (a sonorant) to fill the nucleus slot for a relatively small number of syllables, but in this thesis, I focus on the former type of syllables. The coda is the consonant sequence following the vowel (up to the trailing syllable boundary). An English coda can be a sequence of zero to four postvocalic consonants. Both onset and coda are optional in English, except for the obligatory coda after a (distributionally) lax vowel. The syllable structure of AE can be summarized in the following formula: $C_{0\text{-}3}VC_{0\text{-}4}$ (Abercrombie, 1967: 75; Brown, 2015: 89; Collins et al., 2019: 175).

Many syllable structures can be generated from the schema above. However, the English language imposes certain phonotactic constraints on syllable structures in terms of components. To elaborate on prevocalic consonants, for example, in CV syllables, any consonant can occupy the onset of the syllable except for /ʒ/ and /ŋ/. In the case of a CCV syllable, 26 different consonant clusters are possible, but they all abide by certain phonotactic constraints that divide them into three types. Affricates, voiced fricatives, and semi-vowels /h, j/ are never part of any onset cluster. The lexically most productive CC onset contains initial /s/ (but /ʃ/ before /r/) + plosive or sonorant (e.g., /spɑt/, /stɑp/, /skɪp/, smæk/, /snup, /slop/, /swit/ and /ʃrɛd/). The second type of onset cluster is composed of any obstruent (other than /s, ʃ/ + sonorant /w, l, r/, e.g., /glæd/, /brænd/, and /dwel/. The third type of CC onset is /Cj/, which can only occur before long /u/. Here, the choice of C_1 is fairly unrestricted: it can be a plosive (*pure*, *beauty*, *dune*, *tune*, *cute* but not /gj/), nasal (*mew*, *new*), fricative (*few, view, suit* but not /zj/), liquid /l/ (*lute* but not /rj/) or even a semivowel /h/ (*huge*, but not /wj/). Moreover, the /j/ is dropped in most varieties of American English if C_1 is alveolar.

According to the AE syllable structure schema, English can have up to three consonants in onset (e.g., *strengths* /strɛŋθs/), and as many as four in codas (e.g., *glimpsed* /glɪmpst/). However, there are some phonotactic constraints of AE syllables on this type of cluster as well. In the case of the prevocalic #CCC cluster, the cluster must start with /s/, the second consonant should be essentially a voiceless stop, and the third must be a glide or a liquid.[14] Which consonants can follow each other in syllable onsets and codas is additionally governed by the sonority sequencing principle. This notion as a phonotactic constraint is elaborated more in Subsection 2.2.3 below.

To conclude, and as a universal implicational order, if a longer sequence of Cs is possible, then a shorter sequence is also possible (see subsection 2.2.3. on ISCH below). Thus, if CCC exists, CC exists; if CC exists, then C exists. The longer sequences are always a proper subset

[14] The # symbol stands for the word boundary, which in a monosyllabic word in citation form coincides with the syllable boundary.

of a shorter sequence plus one extra C. In #CCC, the leftmost C must be /s/; in CCCC#, the rightmost C must be /s/.

2.1.3.4. Stress

Stress is defined as the degree of prominence perceived on each syllable, which is caused by specific phonetic manifestations, such as extra loudness, pitch change, longer segment duration, and expanded vowel quality (e.g., Van Heuven, 2018). These characteristics are combined differently depending on the language. The English language tends to combine all four of these components to manifest stress (e.g., Cruttenden, 1997: 13). Additionally, from an articulatory perspective, the syllable-to-syllable variation in stress arises from the changes in the amount of muscular effort the speaker expends in each of the pulmonary, laryngeal, and supralaryngeal stages of speech production (Roach, 1991: 85; Van Heuven, 2018). Perceiving stress is said to be influenced by a number of factors, e.g., increased segment length, loudness, pitch variation, and quality (Fry, 1958; 1965; Van Heuven, 2018), although a syllable can be made prominent even with only one or two factors present.

In relation to syllable structure, stress placement is important because syllables in connected speech may be stressed differently and with different degrees of stress than in isolation. At least one syllable in a word receives secondary stress if it precedes the primary stress in Polysyllabic words; meanwhile, some monosyllabic (mainly function) words are unstressed and subject to vowel reduction in connected speech. It is important to highlight that the unstressed syllables are correlated with some vowel sounds in English that are characterized by centralized positioning and shortness and are articulated almost exclusively within these types of syllables (Wells, 1982: 120). These vowels are known as reduced vowels. They are not included in this research, as they are considered part of the rhythm. Additionally, their exact set differs based on the dialect and the speaker.

2.1.3.5. Vowel reduction and rhythm

This dissertation focuses on the (incorrect) perceptual representation and production of eleven American pure vowels (monophthongs) by Palestinian Arabic learners of English as a foreign language. It does not include the full diphthongs /ai, au, ɔi/. Moreover, it only considers the vowels in stressed syllables. As a consequence of this decision, a number of interference phenomena in PA EFL speech will not be studied, even though such interference may negatively impact the EFL learners' intelligibility and acceptability. Both English and Arabic have been classified as (predominantly) stressed-timed languages. A stress-timed language has a rhythmic structure in which (primary) stressed syllables tend to recur at fixed time intervals, irrespective of the number of unstressed syllables in between stresses (e.g., Abercrombie, 1967). Typologically, stress-timed languages differ from syllable-timed languages in that syllables in the latter type take up approximately equal time intervals so that the timelapse between two stressed syllables increases linearly with the number of unstressed syllables in between the stresses. The stress timing in English falls out as a by-product of phonetic and phonological processes that affect stressed and unstressed syllables in different ways (Dauer, 1983). The first process is called anticipatory shortening, by which the stressed vowel is increasingly shortened as more unstressed syllables separate it from the stress in the next word (Lehiste, 1977; Fowler, 1981). Unstressed syllables in English have relatively simple structures (typically CV, no clusters, preferably no coda), so that unstressed syllables contain fewer segments and therefore take up less time than stressed syllables. Most importantly, unstressed syllables in English cannot contain full vowels. The only vowels that are allowed in the nucleus of an unstressed syllable are schwa and /ɪ/. These are the two shortest vowels in the vowel inventory of English. Schwa (/ə/) is the shortest (and most frequently occurring) English vowel of all. It is articulated in the center of the vowel space with neutral lip rounding. It is generally not considered a phoneme of English because it is in complementary distribution with all full vowels and diphthongs. Full vowels and diphthongs can only occur in stressed syllables but are realized as schwa in unstressed position. Learning to replace full vowels with reduced vowels /ə, ɪ/ is a major difficulty for any EFL learner, even if the learner's L1 is stressed timed. No other language than English is known to have such extreme differences in the articulation of stressed versus unstressed syllables (Dauer, 1983). Although Arabic is listed among the stress-timed languages, there is little difference in the number of segments of stressed and unstressed syllables, due to the relatively simple syllable structure schema of Arabic (see above). Arabic long and short vowels can occur freely in both stressed and unstressed syllables, and – most important of all – Arabic has no extremely short neutral vowel schwa that occurs in unstressed syllables only. EFL learners typically pronounce (their L1 equivalents of the) full vowels in unstressed English syllables, which upsets the stress-timed rhythm and distorts the sound shape of the English words so that word recognition is compromised (e.g., Field, 2005). Since vowel reduction is contingent on the stress patterns and is part of the rhythmic structure of English, we consider it a prosodic phenomenon.

In fact, researchers do not agree on the number of stress degrees in English. Some scholars report four degrees of stress, i.e., primary, secondary, tertiary, and unstressed (Rogers, 2000: 94; Zsiga, 2013: 355). Others maintain that there are only three types, i.e., primary, secondary, and tertiary only (or unstressed, unaccented stressed, or accented stressed), and others propose only two types of stress, i.e., primary and secondary (Jackson, 1980). The rule system proposed by Chomsky and Halle (1968) generates indefinitely many stress levels, with increasingly lower levels of stress as the structures get longer and contain more embedding.

Although there is a great deal of regularity in the English stress system, advice for learners of English as a foreign language is that lexical stress in English is "an idiosyncratic property" (Zsiga, 2020: 150) and "…it is best to treat stress placement as a property of the individual word" (Roach, 1991: 88). In English and similar languages, incorrect placement of word stress is detrimental to correct word recognition (e.g., Field 2005; Cutler, 2015; Van Heuven, 2022).

The rhythm in AE is stress-timed (same as Arabic), i.e., stressed syllables tend to recur at regular temporal intervals irrespective of the number of unstressed syllables in between stresses.[15] However, AE stress is less predictable than Arabic stress because it has many more exceptions to the rules. A word is perceived, among other things, according to its syllable structure as well as its assigned stress in the case of polysyllabic words (Cutler, 2015). Since the position of AE stress is complex and has many exceptions to the rules, learners of AE as L2 will not be able to instantly acquire the stress patterns as they may not correctly perceive them as part of the phonological knowledge even when their L1 is also stress-timed (e.g., Arabic). Field (2005) highlighted that an incorrect stress position compromises word recognition in English both for native and for nonnative listeners. Accordingly, differences in stress patterns even within two stress-timed languages (Arabic & English) may still cause intelligibility problems for AE L2 learners.

2.1.4. Contrastive analysis discussion

In this subsection, I present hypotheses and predictions of different types of possible errors that Arabic L1 learners of English may make, which represent some pedagogical challenges in acquiring L2 proficiency. In light of the previous literature about the two (three) languages, I attempt to explain the dissimilarities between their segmental and suprasegmental features. The participant's L1 background will be considered as a joint mixture of both MSA and PA since only minor differences exist between them in comparison with AE. Additionally, it is difficult to draw a clear line that delineates the exclusive features of these two Arabic varieties. Accordingly, it should also be clear now that the different phonological systems of Arabic and English cut both ways and will affect learners' L2 acquisition, whether their L1 was Arabic or English. However, with the present emphasis I put on AE as L2, I will perform this analysis

[15] In the case of syllable timed languages, all syllables are equally long, whether stressed or not.

from the perspective of PA L1 learners of AE, starting with segmental features (vowels, consonants, and) followed by some suprasegmental differences.

2.1.4.1. Prediction of problems in vowel perception and production

It is well-established that the role of vowels in contrasting accents and dialects of English is larger than that of consonants. Many researchers have stressed this point. Ladefoged and Johnson (2015) state that "Accents of English differ more in their use of vowels than in their use of consonants" (p. 41). Other researchers relate this issue to the fact that vowels are more affected by their sentential position in terms of their duration, intensity, and pitch than consonants (Gleitman & Wanner, 1982; Morgan & Demuth, 1996; Nespor & Vogel, 2008). Accordingly, it can be claimed that if there is a variation in vowels among native speakers of English, it is uncontroversial that they will also vary among nonnative speakers and affect their nonnative speakers' proficiency or even intelligibility.

The allophonic nature of some AE vowels makes PA speakers confuse and even gloss over AE vowels both in perception and production, especially in the case of the mid vowels. Below, I hypothesize the perception and production problems that PA speakers may experience while learning AE vowels. The left-hand part of Figure 2.3 is a schematic representation of the AE monophthongal vowel system. Encircled vowels do not occur in Palestinian Arabic. Vowel enclosed in diamonds are close to one of the PA vowels, either long or short, but saliently deviate from them. The right-hand part of Figure 2.3 charts my hypotheses on how PA vowels are overlayed on AE monophthongs.

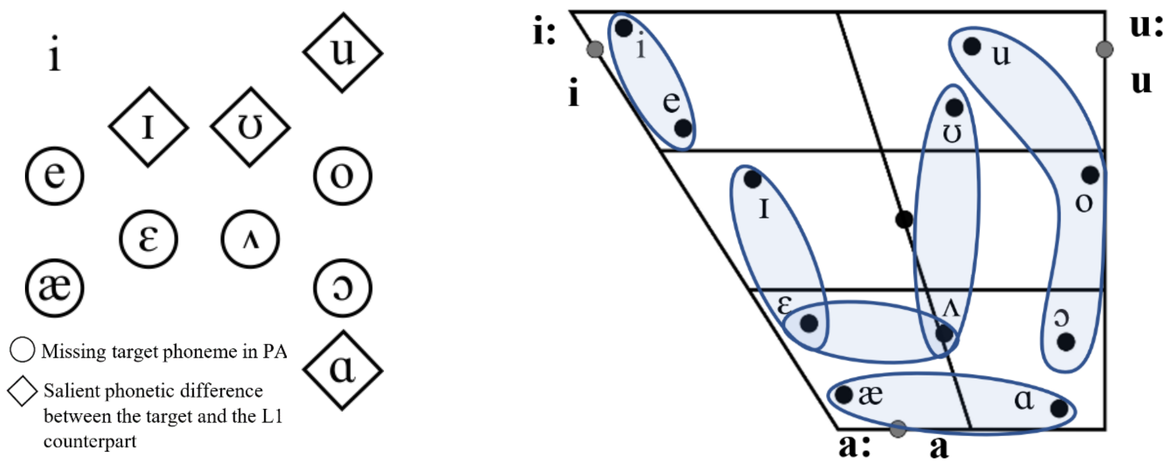

Figure 2.3. PA vowels (on the edges of the IPA quadrilateral) overlaid on AE monophthongs. Vowels enclosed by an ellipse will be insufficiently contrasted by PA learners of American English.

As seen in Figure 2.3, there are almost twice as many AE vowels as MSA vowels. The vowels that exist in the learner's L2 (AE) but not in their L1 are hypothesized to be challenging and difficult to acquire. Flege’s (1995) Speech Learning Model (see section 2.3.2) predicts that these ‘new’ sounds will eventually be incorporated as authentic categories in the learner’s conception of the L2. The vowels in diamonds correspond to Flege’s ‘similar’ vowels. These are predicted to remain a lasting challenge, and will eventually lead to a merged category of the L1 and L2 counterparts, which will sound noticeably wrong to monolingual speakers of both the L1 and of the L2. Table 2.7 below presents my provision of the PA learners of L2 AE errors at the vowel level.

Table 2.7. Provision of PA substitution and learning errors of AE vowels.

AE Vowel	Expected substitution or learning error
i	Mostly realized as PA /i:/ based on duration and tongue position, may be confused with AE /e/ or lax /ɪ/.
e	Absent in PA, confused with AE /i/ as a closest tense counterpart or lax /ɪ/.
ɪ	Mostly realized as PA /i/ and confused with AE lax /ɛ/
ɛ	Absent in PA, may be realized as PA /a/, and confused with AE lax /ɪ/ and /ʌ/ or tense /e/ as a tensed counterpart
æ	Mostly realized as PA /a:/ and confused with AE /ɑ/ based on tongue height
ɑ	Absent in PA. It may be realized as /a:/ based on tongue height and duration
ʊ	Mostly realized as PA /u/ but may be confused with AE /ʌ/
ʌ	Mostly realized as PA /a/, may be confused with AE /ɛ/ and /ʊ/ based on duration and tongue positioning
u	Mostly realized as PA /u:/ based on duration and tongue position, may be confused with AE /o/ and /ɔ/ based on tongue backness
o	Absent in PA, difficult to be recognized, may be confused with AE /u/ or /ɔ/
ɔ	Absent in PA, difficult to be recognized, may be confused with AE /u/, /o/, or /ɑ/

Yavaş (2011: 19) stressed that even if some vowels among different languages are transcribed and labeled the same, there may still exist fundamental differences between them that are audible to the native listener but would escape the self-perception of the L2 learner. This situation can be projected on the free allophones of PA: /e/ and /o/ vowels. The first challenge that faces the L2 learner in such a situation is that the L1 free allophones are not stable in the sense that other point vowels can freely exchange places with them in utterances. The second challenge is that even if they were transcribed with the same symbols that exist in the AE IPA vowel chart, this does not automatically ensure that the same fundamental quality is shared cross-linguistically between L1 and L2.

Ultimately, all vowel contrasts in English will cause a learning problem. On the surface, however, the realization of /i:/ and /u:/ will be reasonable since these vowels will have a sufficiently peripheral quality and will be pronounced long (using the length of the Arabic long/tense counterparts). All other vowels will be problematic and confused with their nearest counterparts, especially for the AE midsection.

2.1.4.2. Prediction of problems in consonant production

On the surface level, many consonant sounds are shared between Arabic and English, but the parameters may differ reasonably to affect the acceptability if not the intelligibility of the spoken sounds. Figure 2.4 below highlights the differences in consonants between PA and AE consonant inventories.

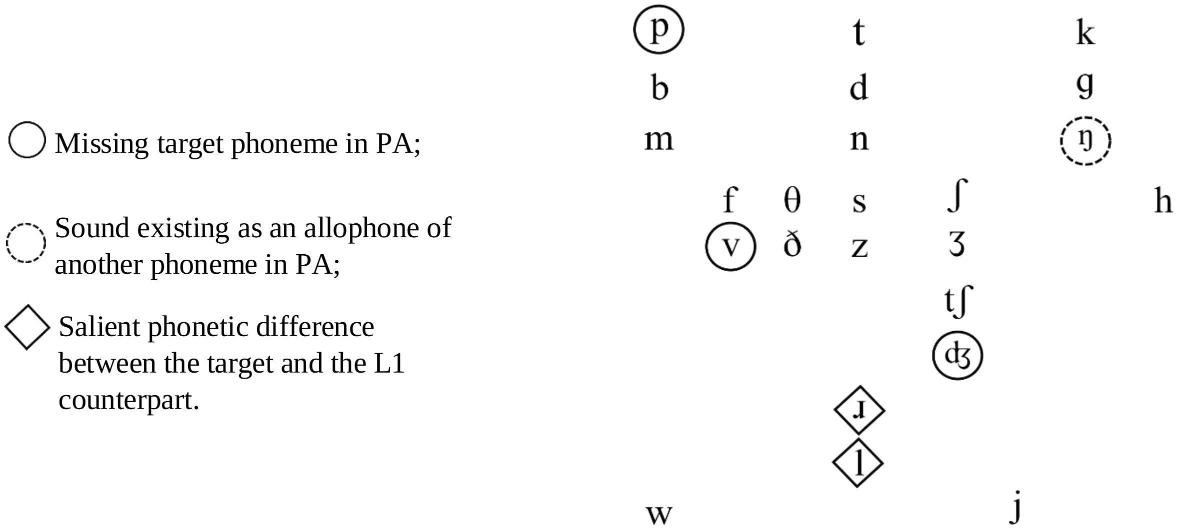

Figure 2.4. PA consonant inventory overlayed on AE consonants. After Yavaş (2011).

In the process of learning new languages, learners may find that many sounds of the new language inventory are shared with their native language sound inventory. Learners may also find some sounds to be identical between the two languages. Based on Figure 2.4, unmarked consonants are predicted to cause no problems at all because they are supposed to be identical and shared between the two languages. Nevertheless, this does not mean that if languages share the same sound symbols, they should have the exact same properties. For instance, the /t/ and /k/ sounds share the same place of articulation, i.e., alveolar in Arabic and English (Elmagdi & Khan, 2005; Alshalaan, 2020). However, it is important to remember that all voiceless plosives (including /p/) are aspirated word-initially in English but not in Arabic. The aspiration of voiceless plosives is strongest in the word-initial position in a stressed syllable. Aspiration is weaker in the word-initial position at the beginning of an unstressed syllable and weaker still word medially at the beginning of a stressed syllable. Aspiration is optional in the word-medial position at the beginning of an unstressed syllable and in the word-final position. Finally, there is no aspiration at all whenever the voiceless stop is preceded by a tautosyllabic /s/ (Balogné & Szentgyörgyi, 2006: 22–23; Ladefoged & Johnson, 2015: 76–77).

Such differences between Arabic and English do not significantly affect the English learning process in the sense that they do not materially hinder communication in English. Nevertheless, these differences are expected to be problematic to varying degrees in relation to the extent to which a native-like perception and production are demanded or expected. On the other hand, there are some consonants in AE that do not exist in Arabic. One example is voiced /v/ and another is voiceless /p/. Many studies (e.g., Ababneh, 2018; Al-Zoubi, 2019; Alshalaan, 2020) report that Arabic L1 speakers are not able to contrast and correctly pronounce sounds such as the /v/ sound (as in /f/ and /v/ pair) and /p/ sound (as in /p/ and /b/ pair) correctly. Table 2.8 lists possible learning errors by PA learners of AE at the consonant level.

Table 2.8. Provision of PA substitution and learning errors of AE consonants.

AE Consonant(s)	Expected PA substitution/learning error
/p/	Absent in PA, confused with PA voiced /b/, expected differences in VOT
/b/	Similar counterpart available in PA
/t, d, k/	Similar counterparts available in PA
/g/	Available in PA Bedouin consonant inventory, may be problematical for urban and rural PA dialects speakers
/m, n/	Both have identical counterparts in PA
/ŋ/	Absent in PA, may be realized as CC with two different places of articulation (voiced alveolar nasal /n/ and voiced velar plosive /g/) broken by a vowel
/f, θ, ð, s, z, ʃ, ʒ, h/	Counterparts available in PA
/v/	Absent in PA. It may be confused with its PA voiceless counterpart /f/
/ʧ/	Possible to be confused with /ʃ/; counterparts available in PA without breaking the CC based on syllable schema
/ʤ/	Absent in urban PA, may be realized as CC with two different places of articulation (voiced alveolar plosive /d/ and voiced palato-alveolar fricative /ʒ/) broken by a vowel
/ɹ/	Absent in PA. it may be realized as alveolar trill /r/
/l/	Available in PA. It may be confused with PA emphatic /lˤ/, but no perceptual confusion will arise from the differences
/w, j/	Both are available in PA, and no confusion is expected
/p^h, t^h, k^h/	Aspirates not available in PA; replaced by non-aspirated counterparts

2.1.4.3. Prediction of problems in consonant cluster production

Based on the stress section discussion for each language, it is clear that the structure of PA and AE syllables differ substantially. The syllable structures allowed by PA constitute a small subset containing the less complex possibilities of AE. PA idiosyncrasies in syllable structure will be transferred by learners to their EFL, and surely learning issues will occur accordingly. Such learning issues will invoke (nonnative) repair strategies (such as inserting epenthetic vowels) whenever they have to pronounce a syllable structure that is outside the intersection of PA and AE (i.e., outside the shared common core). For example, to pronounce consonant clusters word-initially, most colloquial Arabic EFL speakers break up an English CC cluster by inserting an epenthetic vowel before or after the first C. Additionally, CCC clusters are broken up by epenthetic vowels separating each CC sequence. Therefore, words such as *street* or *spring* will be pronounced as /ʔis.ti.ri:t/ or /si.ti.ri:t/ and /ʔis.pi.riŋ/ or /si.pi.riŋ/, respectively.

Similar to Arabic, AE has both open and closed syllables. However, AE allows a syllable to have a two-segment onset cluster, unlike MSA, or three consonantal sounds, unlike PA, and the same for codas with up to four segment consonant clusters. Moreover, AE has null-onset vowel-initial syllables (e.g., *also*, *at*, *or*, *our*, etc.), while Arabic does not allow a zero-onset syllable and must add a glottal stop /ʔ/ (an epenthetic consonant) to overcome this difference in phonological analysis and pronunciation. Moreover, a nucleus-only syllable (e.g., *I*, *eye* /aɪ/, *owe* /oʊ/ or *above* /ə.bʌv/) is allowed in AE but is prohibited in Arabic.

Each language has a particular way of combining sounds to construct words or parts of words within certain combinations and restrictions. Therefore, even when AE allows consonant clusters in onsets and coda, it does so with phonotactic restrictions (see section 2.2.3. on Sonority Principle). For example, there are no words in AE that start with /sd/, /fp/, /bfj/, or /zbf/.[16]

2.1.4.4. Prediction of problems in stress

Many studies show that there can be a transfer from learners' spoken L1 stress patterns to the production of L2 stress patterns in at least two ways: in the positioning of stress at the word level and in the phonetic cues employed in actualizing prosodic contrasts. Such studies examined evidence of prosodic transfer in the production of L2 stress cues. Arabic and AE employ stress to signal the salient syllable at the word level but differ in prosodic patterning. Lexical stress has an important role in native speakers' perception and production of speech. For instance, Arabic learners of AE may face problems with the unpredictable stress patterns of English because lexical stress in AE can alter the meaning of the words, e.g., ˈ*con.trast* (n.) and *con*ˈ*trast* (v.), or ˈ*re.cord* (n.) and *re*ˈ*cord* (v.), which Arabic EFL learners do not have in their L1 (the stressed syllable is indicated by a preceding stress mark: 'ˈ').[17] These and other scenarios will be discussed further below. However, suffice it to say now that stress is essentially predictable in MSA and follows general rules based on syllable patterns that differ from the stress patterns of widely accepted colloquial Arabic varieties (CAVs), e.g., Egyptian Arabic or Eastern varieties, e.g., Levantine Arabic (Ryding, 2005: 36).

[16] For a more elaborated account of English syllable structure, see, e.g., Harley (2006: Ch 3).
[17] By notational convention, the stress mark replaces the syllable boundary symbol '.' in the examples.

Although English stress is rather irregular, it is not the case that there are no regularities. English stress is basically weight-sensitive. However, what makes it difficult for the nonnative learner is that there are higher-order principles that should be applied before the basic rules are applied. For instance, morphologically complex words have to be decomposed into stem morphemes and affixes first. Stem morphemes receive their stress as if they were simplex words, but then the main stresses in one of the stem morphemes must be made stronger than the other main stresses. Which stem in the compound gets elevated depends on the lexical category of the trailing stem: it gets the highest stress if it is an adjective (e.g., *sky*ˈ*blue*, in all other cases the highest stress goes to the leading-edge stem (e.g., ˈ*sky.scra.per*; for classical work in this area see Chomsky & Halle, 1968). Moreover, some suffixes change the stress position in the preceding within-word constituent; these are the stress-attracting suffixes such as *-ity*, which always attract the main stress to the immediately preceding syllable, as in '*com.mune* ~ *com*'*mu.ni.ty*. Last but not least, the sequence of stresses can be adjusted in the postlexical stage by the application of the obligatory contour principle (which dictates that stresses should alternate as much as possible to maximize the number of peaks in a sequence – for two peaks to be distinguished, they must be separated by a valley, e.g., **sky*ˈ*blue* ˈ*ja.cket* > ˈ*sky.blue* ˈ*ja.cket*, see Liberman & Prince, 1977). Because of the intricacy of English stress, EFL learners are advised to simply learn the stress pattern of each word by heart and not bother about regularities.

2.2. Theories and models of L2 perception and production

This section of the chapter is an overview of the theoretical and contextual settings related to speech learning, perception, and production of vowels and consonants, mainly by nonnative learners. The section is divided into three sections. This section (2.2) reviews the following three specific hypotheses. Namely, the Contrastive Analysis Hypothesis (CAH, Lado, 1957; Wardhaugh, 1970), the Markedness Differential Hypothesis (MDH, Eckman, 1977), and the Interlanguage Structural Conformity Hypothesis (ISCH, Eckman, 1991), all of which are employed to assist in analyzing L2 phonology and to account for L2 acquisition difficulties in EFL settings that are relevant for this thesis.

In section (2.3), the aim is not to discuss all models related to L2 learning but to focus on specific, the most recent and dominant models in this field of study. Therefore, this thesis is based on an in-depth review of three theoretical models, as its primary aim is to tackle the influence of L1 phonological inventory on L2 perception and production as well as on the ultimate attainment of L2 speech proficiency or the lack thereof. The three relevant L2 speech perception & production models are the Perceptual Assimilation Model (PAM) and its extension to the Second Language (PAM-L2), the Speech Learning Model (SLM), and the Second Language Linguistic Perception model (L2LP). In terms of L2 learning, these three models make opposing predictions, which can be tested against the empirical data provided by the present study.

2.2.1. Contrastive Analysis Hypothesis (CAH)

To create (or increase) cross-language awareness, contrastive linguistics has a very important part to play (Lado, 1957; Mair, 2005; Wardhaugh,1970, 1974). CAH (Lado, 1957, 1964) explains that L2 speakers generally depend on their L1 rules, either to fill the knowledge gap or to compensate for their incorrect perception of the L2 sound structures, which consequently leads to different errors in pronouncing or forming the utterances. Many studies claim that such errors can be avoided through the repetition of the correct forms and pronunciations until the errors are fixed (Fries, 1945; Gathercole, 2006; Gathercole et al., 1992; Pallier et al., 1999). Following this train of thought, errors were considered a sign of each learner's difficulties as a symptom of transfer (or generalizing) habits from L1 to L2. Lado (1957) states that language users tend to transfer the meanings, forms, and their distribution of their native language and culture to foreign languages and cultures in two reciprocal ways: productively, when practicing speaking the language and acting in the culture, and receptively, by trying to grasp and understand the foreign language and its culture as practiced by the natives (Lado, 1957: 2).

Simply put, CAH (Lado, 1957) states that learners depend on their culture and L1 in transferring forms, meanings, and their distributions to L2, both in perception and production, in what is termed the transfer phenomenon. In addition, the structural differences between the two languages being learned are imperative and good enough to account for L2 learning. Accordingly, Lado (1957) further discusses that a strong form of the CAH is established when L2 phonology is filtered through the learners' L1 leading to two types of interaction. When similarities are found (or predicted) between L1 and L2 forms and patterns (of phonologies), a positive transfer is expected from L1 to L2. However, differences between the L1 and L2 phonologies will lead to difficulty in learning these different features, leading to what is called a negative transfer (i.e., interference). Therefore, learners will depend on the already existing features in L1 to compensate for the missing (or difficult) features of L2 in order to perceive and produce them more readily.

Another notion originally related to CAH is the 'interlanguage' system, which is basically a version of a language acquired as L2 that is shaped by different factors, such as language transfer, learning strategies, overgeneralizations, etc. It is argued that the linguistic system acquired by a learner as a second language is different from both of the learner's L1 and L2, but linked nonetheless (Tarone, 1988). Selinker (1972) highlighted that in the L2 acquisition process, learners attempt to produce native-like forms that result in developing their own L2 linguistic system, yielding a new system that differs from their L1 and L2. This system is mostly referred to as the interlanguage (IL) or what Nemser (1971) called 'an approximative system' or as Corder (1971) calls it 'an idiosyncratic dialect'.

Furthermore, the IL was developed to account for the language system of adult L2 learners, which can be defined, from another perspective, as an L2 system that mainly includes learners' own productions of a particular language's output during the development process of their linguistic system (Major, 2014; Selinker, 1972). Thus, many components interact to formulate the learner's IL. More specifically, the elements from L1 (based on transfer phenomena), elements from L2, and universal elements (or implications) that are missing from both L1 and L2 all together compile the interlanguage system (Corder, 1971, 1981; Major, 2001), which allows IL to be extended to other hypotheses related to 'markedness' that will be discussed below.

The viability of the CAH has been the subject of controversy, as it was claimed to explain and predict all learning errors based on transfer. There are two forms of CAH, viz. the strong and weak forms, which are suggested to minimize the role of negative transfer at different levels of Second Language Acquisition (SLA). Moreover, several research studies have reported that the transfer phenomenon itself at the phonological level is salient and obvious in SLA, and the CAH explanation for the interference phenomenon cannot be easily denied (Brown, 1994; Selinker, 1992). The strong form of CAH has been found to make some wrong predictions about the L1-L2 interaction, which led to the emergence of a weak version of CAH by Wardhaugh (1974), which states that a linguist should "…use the best linguistic knowledge available to him in order to account for observed difficulties in second-language learning." (p. 81). There is one problem with the weak version, which is limiting the process to the observed difficulties and not predicting possible errors that will be made in L2 acquisition. Therefore, Eckman (1977) argued that the weak version cannot function as a principle of the CAH for L2 acquisition since it cannot be falsified. Hence, I adhere to Lado's (1957) strong version of the CAH in this research, as it provides successful explanations for many transfer phenomena at the phonological level.

Briefly, the CAH, at the phonological level, stresses that pronunciation errors are considered possible to be fixed with more drilling while focusing on the problematic L2 sound patterns.

Moreover, as the errors are the result of interference (through negative transfer), the learning difficulties can be predicted by initiating a structural comparison of the L1 and L2 sound inventories (e.g., Lado, 1957, 1964; Wardhaugh, 1970; Broselow, 1984). The belief that pronunciation errors are linked in origin to L1 interference in L2 led to the rise of Contrastive Analysis (CA) in the field of phonology as a theory highly influenced by behaviorism.

In relation to CAH and IL, another two hypotheses were proposed as a revision of CAH. Namely, the Markedness Differential Hypothesis and the Interlanguage Structural Conformity Hypothesis. These incorporate the notion of markedness and can be employed in analyzing L2 phonology and account for L2 acquisition difficulties based on the learner's IL.

2.2.2. The Markedness Differential Hypothesis

Markedness Theory first emerged as an idea of markedness that dates back to the Russian linguist Trubetskoy (1939) and was formalized later by Chomsky & Halle (1968: Ch 9). The MDH was presented as an alternative to the CAH by Eckman (1977), and it is substantially rooted in theories of markedness to measure (among other disciplines) the relative phonological complexity (Colantoni et al., 2015). Eckman's (1977) MDH is, in fact, a reformulation of CAH that includes the concept of typological markedness (Eckman, 2008). More specifically, MDH was proposed to address some empirical issues of CAH (Eckman, 2012), especially for L1-L2 being necessary but not enough to address L2 learning problems; therefore, typological markedness (degree of difficulty) should be employed as an explanation of such difficulties (Eckman, 2008). According to the MDH, the more implicationally marked an L1 form is, the more difficult its acquisition will be in comparison with those forms that are less marked or considered universal implications, i.e., when they are widely distributed across languages (Colantoni et al., 2015: 64; Eckman, 1977: 320).

Eckman (1977: 320) defined Markedness as follows: "A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A." This, in essence, depends on how general or common a given structure/feature is across the world's languages. Within MDH, markedness is delineated as areas of difficulty that can be predicted based on a systematic comparison of the grammars of the native language, the target language, and the markedness relations stated in universal grammar based on three principles (Eckman, 1977: 321; Eckman, 2012: 410):

1. Those areas of the TL that are more marked than the native language will be difficult.
2. The relative degree of difficulty will correspond to the relative degree of markedness.
3. Those areas of the TL that are not more marked than in the NL, will not be difficult.

The idea behind the concept of markedness is that the binary oppositions between specific linguistic attributes cannot be simply taken as polar opposites. The criteria for establishing the marked member of an opposition pair of sounds, for instance (or binary feature), are based on an implicational relationship that states the occurrence of the marked member may it be a sound or a structure (e.g., complex CC onsets, or voiced obstruents) implies the occurrence of an unmarked member (e.g., singleton C onsets, or voiceless obstruents) but not *vice versa*, and this assumption holds true both across languages and within one single language, irrespective of context or circumstances (Eckman, 2012: 410). To elaborate with examples related to the languages addressed in this thesis, the occurrence of voiced stops implies the occurrence of voiceless stops. This means the /p/ sound, for example, is unmarked and the voiced version of it /b/ is more marked, i.e., more difficult, or the /v/ sound (being marked) implies the occurrence of the /f/ sound. As per this hypothesis, the occurrence of the (marked) /b/ sound in Arabic will make it easy (on the surface level) for Arabic learners to acquire the (unmarked) /p/ sound of English but not the other way around. However, the aspiration of English /p/ and its relation to different stress patterns bear difficulty for Arabic learners and will be marked for them as they lack such features in their L1. This contrast set is particularly important because many Arabic L1 studies on L2 English pronunciation have reported incorrect perception and/or production of these specific sounds (Farrah & Halahlah, 2020; Jabali & Abuzaid, 2017, on L1 Palestinian Arabic; Hassan, 2014, on Sudanese Arabic; Ababneh, 2018, on Saudi Arabic; Kharma, & Hajjaj, 1989, on different Arabic varieties), as many L1 Arabic learners face difficulty in discriminating between the two sounds and in producing the voiceless counterpart /p/, precisely because of the marked nature of English [p^h]. On the contrary, Eckman (1977) explicitly stated that “there are apparently no languages with just voiced obstruent phonemes” (p. 320, last paragraph), claiming that the occurrence of voiced obstruent phonemes is more marked than their voiceless counterparts, which are less marked. Interestingly, Arabic has the /b/ sound but does not have a /p/ sound.

In the case of Arabic voiced obstruents, the first principle (out of the three mentioned above) of MDH (Eckman, 1977) applies, but the third part does not. Therefore, AE /p/ sound is supposed to be more difficult for PA learners of AE as L2, but not the other way around.

The MDH assumptions concerning other markedness relationships for Arabic fricatives are sustained, especially with the occurrence of /f/ without the occurrence of /v/. Therefore, it is presumed that it would be easy for the PA participants to pronounce the AE /v/ sound effortlessly. The presence of the marked voicing contrast in the pharyngealized consonants, such as velar/uvular fricatives /χ/ and /ɣ/ in Arabic, which is less frequent (i.e., less common and, by default, more difficult) among the world languages, makes it easier for PA EFL learners to acquire the less marked voiced labial fricatives /v/ (more common on a global scale). [18]

In summary, CAH tries to analyze L2 learning difficulty based on the attributed differences from a binary scale between L1 and L2, while MDH refines this based on the degree of markedness by stating that learning difficulties arise only if the L2 has a marked member of an opposition for which the L1 only uses the unmarked possibility. Accordingly, the MDH confirms that the marked structures within the areas of differences between L1 and L2 are more difficult to acquire than the related unmarked forms. MDH further details that not all L1-L2 differences will have resulted in equal difficulty. This means that if the L2 structures are not related by markedness principles to any L1 structure, then they are predicted to cause no difficulty or learning problem even if they are different than in the L1. The opposite is true when marked L2 constructions constitute a learning problem, and the level of difficulty involved will be related to the degree of markedness.

In support of the MDH, some pieces of evidence are provided to highlight that the problems that occur in learners' perception and production cannot be explained based on just the presence of differences between L1 and L2 but that the typological markedness predicts (and explains) the difficulties that learners encounter. In this relation, what is known as "directionality of difficulty" is present in cases where speakers of different L1 backgrounds start learning the same L2, but some of them experience more difficulties than others.

2.2.3. The Interlanguage Structural Conformity Hypothesis

The validity of MDH has been shown in a number of studies (e.g., Anderson, 1987; Carlisle, 1991 among many others) backed by errors found in IL based on L1 transfer, but it mainly compared dichotomous (marked vs. unmarked) structures without predicting on certain structures that exist in both L1 and L2. Eckman (1991, 1996, 2008) considered this as a shortcoming in MDH and consequently proposed the Interlanguage Structural Conformity Hypothesis (SCH), which (like the previous hypotheses) involves typological markedness and IL but isolates the effects of L1-L2 differences from the equation. Furthermore, many of the studies in IL phonology have focused on the impact of implicational universals on L2 phonology acquisition since the publication of MDH (Anderson, 1987; Eckman, 1987, 1991; Eckman & Iverson, 1993; Major, 1990; Major & Faudree, 1996).

[18] Only 4.1% of the surveyed languages in the WALS contain pharyngealized consonants (Maddieson, 2013c), while 91.3% of the languages have fricative consonants and that only 8.7% of the languages do not have them (Maddieson, 2013d).

Eckman (1996) argues that the primary motivation behind ISCH is that an L2 pattern can be an error pattern even if the structures at hand are not related to L1-L2 differences (i.e., surfaces in the L2 acquisition process in only an IL stage). This is a case not explained by the earlier MDH. Hence, a solution can stem from eliminating the L1-L2 differences as a criterion for invoking markedness in explaining L2 acquisition issues. Therefore, it can be argued that if learners will perform better with less marked structures relative to marked structures, then the MDH becomes a special case of ISCH where the universal implication (or generalization) holds true for the interlanguage, and these structures, supported by universal implications (or generalizations), are the ones that differ between L1 and L2. The ISCH hypothesis confirms that there are great similarities between interlanguages and primary languages in at least one primary aspect, which is that they both obey the same set of implicational universals. Concerning ISCH, Eckman (1991) stated that "The universal generalizations that hold for primary languages hold also for interlanguages" (p. 24).

The ISCH can be employed to explicate the difficulty and progression of acquiring EFL consonant clusters. In this sense, I discuss the following two implicational universals: the Fricative-Stop Principle and the Resolvability Principle. The Fricative-Stop Principle (Eckman, 1991) states that "If a language has at least one final consonant sequence consisting of stop+ stop, it also has at least one final sequence consisting of fricative+stop" (p. 24). The Fricative-Stop Principle denies the possibility for a language to have words that end with a stop+stop consonant sequence without also having words that end with fricative+stop clusters. At the same time, it claims that a language can contain words that end with fricative+stop clusters while not necessarily permitting a stop+stop cluster word-finally. Therefore, a case of stop+stop cluster word-finally is more marked than a fricative+stop cluster word-finally. Meanwhile, according to the Resolvability Principle as a universal implication, if a language contains a consonant cluster of length *x* in either the beginning or final position, it also has at least one continuous subsequence of length *x*–1 in the same location (Eckman, 1991; Greenberg, 1978; Kaye & Lowenstamm, 1981). In essence, Vennemann (2011) stated that implicational universals involve at least two properties in a conditional relationship. If X then Y, e.g., if a language has a CCC onset, then it must have a CC onset.[19]

Vennemann (2011) presents a number of preference laws for syllable structure that specify the preferred syllabic patterns of natural languages and determine the direction of the syllable structure change. Part (a) of the syllable Head (or onset) Law (Vennemann, 2011) states that

[19] Greenberg points out that this IL principle is valid only as long as the syllable is not CV, which does not necessarily entail the presence of V onset-free (i.e., zero onset or at least nucleus-only syllable).

"A syllable head is the more preferred: the closer the number of speech sounds in the head is to one" (p. 13). Vennemann (2011: 14) also mentions that "For heads with more than one speech sound, part (a) of the Head Law suggests changes reducing their number." Furthermore, Part (a) of Vennemann's (2011: 21) Coda Law stipulates that "A syllable coda is the more preferred: (a) the smaller the number of speech sounds in the coda." From these two laws, it can be inferred that the most preferred syllable is the $C_1V_1C_0$ syllable, where a single onset C is the optimal onset and null coda is the optimal coda. In relation to typological markedness, this understanding of preference laws by Vennemann (2011) renders CV the most preferable syllable (the most unmarked syllable type) in languages and suggests that an increase in either margin's length of a syllable results in an equivalent proportionate increase in syllable markedness.[20] To exemplify, the pronunciation of the open monosyllabic word *knee* /ni/ is less marked than the closed monosyllabic *kneel* /nil/ or *knees* /niz/, which are less marked than the CC-closed monosyllabic *kneeled* /nild/.

In relation to this discussion, it should be considered how the consonant clusters in onsets and codas (and in syllables as a whole) are subject to phonotactic constraints such as the Sonority Sequence Principle (SSP) or hierarchy of speech sounds. The degree of sonority of a speech sound relates to many factors, such as its loudness compared to other sounds, how much it can be prolonged, and the degree of stricture in the vocal tract. The more open the mouth and the louder the sound is, the more sonorant it will be. Another factor related to this principle is voicing in the voiced/voiceless contrast. The voiced sound is said to be more sonorant than its voiceless counterpart within one language inventory. Moreover, the formant patterns are related to sonority in acoustics; the more sonorant a sound is, the clearer and more distinct its formant structure (Gussenhoven & Jacobs, 2011). The hierarchy of sound sonority is discussed differently by scholars, but the differences are uncontroversial in essence in terms of ordering and only vary in the details of manifesting the hierarchy. Hogg and McCully (1987), as cited in Yavaş (2011), presented a 10-point sonority scale, as shown in Table 2.9 below, where 10 means most sonorant.

[20] Since the maximum onset in language is CCC and the maximum coda can be CCCC, it makes sense to predict that increasing the number of Cs in the coda adds a little less to the complexity/markedness of a syllable than adding elements to the onset.

Table 2.9. Sonority scale of speech sounds. Based on Hogg and McCully (1987)

Sonority values	Sounds	Examples
1	Voiceless stops	/p, t, k/
2	Voiced stops	/b, d, g/
3	Voiceless fricatives	/f, θ, s/
4	Voiced fricatives	/v, ð, z/
5	Nasals	/m, n, ŋ/
6	Laterals	/l/
7	Flaps	/r/
8	High vowels (and glides)	/i, u/
9	Mid vowels	/e, o/
10	Low vowels	/ɑ, æ/

Note. For a more in-depth discussion of sonority scale ordering, see Zsiga (2020: Chapter 7)

At the syllable level, the sonority principle presents explanations of the similar norms of consonant sequencing structure across different languages. Here, I recall the syllable structure of both Arabic and English and the phonotactic constraints thereof. While PA allows for some onset-CC and final-CC under certain conditions and rules, the rules and constraints upon English syllable structure were also given. Gussenhoven and Jacobs (2011) proposed a Sonority Profile that stipulates that "The sonority of syllables increases from the beginning of the syllable onwards, and decreases from the beginning of the peak onwards." (p. 165).

In the case of differences between L1 and L2, or any deviation from the target language norms, or reversing the sonority hierarchy, this sonority principle will posit typologically more markedness weight in pronouncing these deviations.

As a concluding remark for the three hypotheses − CAH, MDH, and SCH − the main reason for L1 interference errors is presumably caused by a lack of contrast between specific L2 sounds and elements in the L1 phoneme inventory. Moreover, many models and theories of SLA focus on the linguistic experience of the individuals to predict and discover how nonnative learners perceive and produce particular sounds of L2 and whether such experience posits any constraints on the speaking process, especially on those sounds/phonemes that deviate from L1, phonologically and/or phonetically (Best & Tyler, 2007). In Subsection 2.3.4, I will provide my understanding of the problems from the perspective of each hypothesis and model, but first, I will discuss three dominant models in the field of language acquisition.

2.3. SLA models

In this thesis, various models are introduced to reveal (among other things) how the two selected languages contrast their segments (vowels and consonants) over the articulatory-acoustic space, as well as their suprasegmental features. Therefore, I analyze and compare the phonotactic features of PA-accented English in light of related theories that can account for speech perception and production of English as L2. Among these models, I will discuss the Perceptual Assimilation Model (PAM; Best, 1994; 1995) and its extension to L2 learning (PAM-L2; Best & Tyler, 2007) as the leading framework for the whole dissertation, the Speech Learning Model (SLM; Flege, 1991, 1995) and its revised version SLM-r (Flege & Bohn, 2021), and the Second Language Linguistic Perception model (L2LP; Escudero, 2005, 2009) and its revised version (van Leussen & Escudero, 2015), all of which aim to explain how L1 experience influences L2 sound learning in terms of learning scenarios.

Although each model focuses on one phase or more of the L2 acquisition level, they all share the assumption that nonnative learners of an L2 often find it difficult to understand native L2 speakers and struggle even more to produce native-like (connected) spoken utterances at the very beginning of the L2 acquisition process (naïve level), and often even after a long time of exposure to L2. Among the reasons for such hurdles is that L2 learners depend on their already developed L1 forms and patterns of speech when trying to produce and perceive L2 sounds (the transfer phenomenon mentioned earlier). Learners in this situation are observed to have a different accent (which may vary from light to heavy) rather than the standard pronunciation of the target language. Several studies (reviewed below) have discussed the psychological and sociological dimensions of having an accent in general, especially for L2 learners.

The results of L1 remnants on L2 fluency (or accent) create new problems for L2 learners in addition to what is mentioned earlier about IL. Many Arabic learners of EFL speak English with an accent that makes them difficult to understand for native English listeners (Farrah & Halahlah, 2020; Jabali & Abuzaid, 2017; Rehman et al., 2020). The researchers suggest that this is primarily caused by a lack of practice, either in a controlled context such as in-classroom practice or even outside the school environment. PA learners are no exception to this generalization. Farran et al. (2020) found that PA learners of EFL lack sufficient oral/aural practice and assessment of EFL inside the schools in a study assessing the English language school-leaving exam in the General Certificate of Secondary Education (GSCE) test for the year 2018. Such situations will eventually affect learners' language proficiency, higher education, profession, and social interactions.

All of the three L2 speech learning models claim that predictions can be yielded about L2 acquisition based on the similarities and differences of sounds in the involved languages. However, the way each model addresses L2 sound acquisition difficulties differs from one model to another, as detailed below for each model. The next section starts the discussion with, arguably, the most widely tested theory of nonnative and L2 speech perception, i.e., The Perceptual Assimilation Model - L2.

2.3.1. The Perceptual Assimilation Model (PAM)

The Perceptual Assimilation Model (PAM, Best, 1994, 1995) and its extension PAM–L2 (Best & Tyler, 2007) address the nonnative perception of sounds by learners who are still at the beginning of their acquisition process of L2 (nonnative listeners). It simply states that second language (L2) learners' difficulties in acquiring L2 speech are governed by their perceptual limitations.

PAM-L2 posits that L2 listeners categorize L2 sound contrasts based on the degree of similarity and discrepancy observed between native and nonnative sounds. More specifically, PAM-L2 predicts how learners will perceive and classify nonnative phones, i.e., the ease of auditory discrimination between two arbitrary nonnative sounds as a contrasting set. These classifications of L2 contrast sets govern how the differences are assimilated to native categories. The perceived L2 sounds (or phones) are assimilated to (i.e., identified, classified, and organized in terms of) L1 sounds differently (speech vs. nonspeech, and in the case of speech, native vs. nonnative) by naïve L2 listeners (or L2 beginners). PAM specifies three main assimilation patterns as follows:

- The nonnative sound is labeled **Categorized** (C) if it clearly assimilates to a single L1 segmental category only. In this case, the assimilated phone is heard either as an identical, an acceptable, or a deviant exemplar of that L1 category.
- The nonnative sound is **Uncategorized** (U) if it is perceived to be within the L1 phonological space but does not clearly fit one specific L1 phonological category. Rather, it falls in between two (or even more) L1 categories for the L2 learners.
- The nonnative sound type is **Non-assimilable** (N) if it does not assimilate into any L1 category and therefore is outside the native phonological space. These sounds are often heard as nonspeech sounds, such as other types of sounds made by the vocal tract, e.g., snaps or clicks, or the flicking of fingers or clapping hands.

From the PAM perspective, L2 learners depend on their L1 phonemic system to perceptually assimilate L2 phones to the most articulatorily-similar native phoneme, which may or may not yield learning problems. Thus, PAM provides predictions about the ease of auditory discrimination between L2 sounds in light of the learner's L1. The classification is based, first, on the previously mentioned tripartite assimilation process. Then, PAM sets out a typology of six types of assimilations (listed below) for categorizing L2 sounds (pairwise) according to their phonological and phonetic distance to L1 phonological categories and accordingly predicts their level of difficulty.

- Two-Category assimilation scenario (TC Type). Two nonnative phonemes are assimilated to two different L1 phonemes, which results in easy and excellent discrimination. In theory, this contrast should not present any learning problem for L2 learners because the two nonnative sounds would not cause any perceptual conflation from L1.
- Category-Goodness scenario (CG Type). Two nonnative sounds are assimilated to the same L1 sound, but listeners are expected to discriminate fairly well between the two sounds based on their goodness-of-fit to the L1 phoneme because their goodness-of-fit degree varies (i.e., one of the two nonnative sounds is a clearly better exemplar of the L1 sound than the other). This scenario predicts a temporal learning problem because eventually a new category boundary is expected for each of the nonnative sounds**.**
- Single-Category assimilation scenario (SC Type). This type occurs when both phonemes of a nonnative contrast are assimilated to the same L1 sound and both L2 sounds are perceived as equally acceptable exemplars of the L1 category (i.e., with an equal goodness-of-fit). Discrimination for this type of assimilation is more difficult and is expected to be poor because the scenario causes perceptual conflation of the two sounds. Thus, long-lasting learning problems are predicted.

Succinctly, the TC contrast is simpler to discriminate than the CG contrast, which in turn is easier than the SC contrast. The classification of L2 contrast sets determines how learners categorize these sets into assimilation scenarios based on their L1 sounds and then implicitly predict either easy acquisition or temporal or lasting acquisition problems according to the level of difficulty in discriminating these contrasting sets.

I only summarized the three scenarios that were found in my own results, which all involve categorization of both L2 sounds to L1 phoneme(s), but PAM also describes another three possible cases for uncategorized sounds that are less likely to emerge in L2-L1 perceptual assimilation in this thesis. In such cases, one or both L2 sounds may fail to sufficiently assimilate to any existing L1 phoneme (one or both are Uncategorized). These patterns are (4) both Uncategorizable (UU Type), (5) Uncategorized versus Categorized (UC Type), and (6) Non-Assimilable (NA Type). For more information on PAM scenarios, I refer to Best & Tylor (2007) and Tylor (2019)

This thesis compares two languages with different sizes of consonant and vowel inventories. Arabic, in general, has a small vowel inventory in comparison to English. Therefore, learners will have relatively few vowel categories to map English vowels onto. For example, /ʌ/ and /ʊ/ are both predicted to map onto Arabic back high short /u/ (an example of a Single Category assimilation pattern).

As for consonants in the PAM framework, for instance, a two-category scenario is predicted for the AE, /t/ and /d/ which have equivalents in Arabic, i.e., /t/ and /d/. However, the AE /v/ phoneme does not have a clear match in contrast with /f/ in PA or MSA. That is, AE /f/ would be a good exemplar of PA /f/ while AE /v/ is also categorized as /f/ but with poorer goodness-of-fit. The same scenario is hypothesized for the AE /k/~/g/ sounds assimilating to Arabic L1 /k/ which also lend themselves to the comparison based on Eckman's (1977) MDH hypothesis mentioned earlier. Therefore, these sounds cause a category goodness scenario. A good example of the SC scenario between the two languages is the AE /b/ ~ /p/ contrast for the Arabic /b/ phoneme, since neither /v/ nor /p/ exist in Arabic (as discussed earlier); they both should fall in the same categorization in their contrast sets, but the AE /p/ sound imposes more difficulty on the ground of its aspirated variation.

Another way PAM considers the perceptual assimilation of L2 sounds to L1 is by categorizing individual nonnative sounds instead of categorizing them into scenarios. To do justice to the relative independence of agreement and typicality PAM-L2, some studies (Guion et al., 2000; Best & Tyler, 2007) suggest that the information provided by agreement and typicality rating be combined into a single measure, which is called the *Fit Index* (a predictor of how well a foreign sound fits a native category).

Guion et al. (2000) proposed quantitative criteria for the classification of nonnative phones using the fit index. The fit index of a nonnative sound is found by multiplying the identification score of the L2 sound as a token of a specific L1 sound in a perceptual identification task by its goodness rating as a token of the L1 category. Depending on the fit index, the L2 sound is then categorized as a good, fair, or poor token. The boundaries separating the three fit categories are set at 1, 0 and −1 SD away from the arithmetic mean of the modal fit indexes for the L2 sounds. This method has been applied to an array of different language pairs (Guion et al., 2000; Park & de Jong 2008; Wang & Chen, 2020), including English learners of Arabic (e.g., Lababidi & Park, 2017 on Arabic consonants). To the best of my knowledge, no previous studies have approached L1 Arabic learners of EFL using this method.

2.3.2. The Speech Learning Model (SLM)

The Speech Learning Model (SLM, Flege, 1995, 2002) and its revised version (SLM-r, Flege & Bohn, 2021) aim primarily to examine the extent to which individuals learn to accurately produce and perceive phonological segments (i.e., vowels and consonants). SLM focuses on the ultimate attainment of the sounds, while the revised version (SLM-r) calls for the necessity of examining the early stages of L2 speech development to comprehend the process of forming L2 phonetic categories (Flege & Bohn, 2021). SLM postulates many tenets to account for L2 speech. Among these tenets, SLM states that L2 learners have constant access to the same basic learning mental faculty throughout their lives (Guion et al., 2000). Additionally, SLM states that L1 and L2 sounds exist in the "same phonological space" (Flege, 1995: 239) (which has been redressed in the latest version of SLM-r as "common phonetic space") and that learners can maintain a cross-linguistic distinction between them (Flege & Bohn, 2021: 21).

Based on SLM, the likelihood of learning new phonetic categories in the L2 can be predicted, both in production and perception. While it takes into account the effect of several sociolinguistic factors on speech perception and production, such as age of arrival in the L2-speaking country, age of L2 acquisition, and L1/L2 use (Flege, 2002), SLM primarily focuses on the attributes of sound segments and argues that L2 phones are "equivalence-classified", where the perception of an L2 sound is constantly assimilated to the nearest L1 phonetic category across the life span and is perceptually related to allophonic "equivalence-classified" positioning within the same phonological space (Flege, 1995; Flege & Bohn, 2021). This equivalence classification process is reciprocal cross-linguistically, as it can enhance the perception of phonetic variability in L1 speech. However, the incorrect equivalence classification performance in L2 can cause problems with the production and/or perception of L2 segments. Therefore, SLM divides L2 sounds into three categories relative to the L1 inventory.

- The 'identical' sounds are the least difficult to learn since they are identical to their nearest L1 sounds in all important qualities. Accordingly, positive transfer of the L1 sound to the L2 is expected to achieve highly accurate L2 sounds. Both L1 and L2 'identical' sounds would be written with the same IPA symbol (Flege, 1995). For example, /d, b, t, j, f, z, ʃ, h, l, m, n, w, θ, ð/ are supposed to be identical in English and Arabic.
- In contrast, L2 phones that are perceived as 'new' sounds will eventually be fully acquired assuming that learners will spot the phonetic differences (or dissimilarities) between the already existing L1 phonemes and the new L2 phone, but still with more difficulty than

'identical' sounds because the new sounds resist equivalence classification with L1 sounds directly, and this process demands more effort to approximate accent-free production and correct perception. To exemplify, I mentioned earlier in Subsection 2.2.1 that there are 'new' sounds in AE, e.g., /ŋ, ʧ, dʒ/, which do not have clear equivalents in Arabic, and they are represented with symbols that do not exist in the Arabic consonant inventory.

- On the other hand, if much phonetic dissimilarity is spotted between an L2 phone and its nearest L1 phone, the production of the L2 phone can be more accurate and easier to grasp than between L1-L2 sounds with less phonetic dissimilarity, i.e., 'similar' sounds. Similar sounds then are transcribed with the same IPA base symbols but differ in the diacritics (in the case of vowels, for example). In the PA case, sounds like /v, p/ have similar voiced/voiceless counterparts. Additionally, AE /l, r/ have corresponding sounds that differ in renditions in Arabic, e.g., [r] in Arabic vs. [ɹ] in AE.[21]

The 'similar' sounds, in theory, are the most difficult for L2 learners to learn since they are similar, but not exactly identical, to an L1 category. Therefore, L2 speakers perceive L2 sounds as good instances of already existing L1 categories, as hypothesized by Flege (1981). This equivalence classification will prevent adult L2 learners from forming a new phonetic category for 'similar' sounds, in contradistinction to the 'new' L2 sounds. SLM, for example, predicts that a new phonetic category will emerge for a certain vowel if this vowel is perceived by the learner to be further away than some threshold distance from the closest vowels in L1. However, for 'similar' sounds, such a threshold will not be exceeded, specifically for adults, so that the learner enlarges the existing L1 category to include the nonnative (foreign-accented) realization of the L2 sound as the phone falls in a merged category between L1 and L2. This results in producing incorrect or poor versions of the sound both in L2 and L1. To overcome such errors, the acquisition process evolves to keep the original category of the L1 phoneme as an optimal but narrower representative (a subcategory) of the L1 phoneme and set up another subcategory for the similar L2 phoneme that is optimal for the L2 tokens – which can be done through long memory activation. Examples of 'similar' sounds are /t/ or /d/ which are dental in Arabic but alveolar in English, or more problematic sounds such as English /p/ perceived and produced as Arabic /b/ (cf. Eckman, 1977 MDH). The differences may not be salient to Arabic learners of English, but their accented production will affect the intelligibility of their speech, or at least, such differences will affect the learner's proficiency, i.e., perceived foreignness.

[21] Here, the heuristic that a difference in an IPA base symbol yields a NEW sound does not work. The variants of /r/ are not indicated by diacritics but by different base symbols.

Additionally, for the SLM, phonetic categories are not static but constantly changing, depending on L1 and L2 experience and usage (Flege, 2003). For instance, in the early stages of L2 learning, SLM predicts that L2 vowels that are rated as phonetically similar to an L1 vowel will be produced fairly well, as in the case of vowels, e.g., /i, ɪ, æ, u, ʊ/. Additionally, SLM predicts that L2 vowels that are rated dissimilar to any L1 vowels will be poorly produced and may be confused with one or more L1 vowels, e.g., /e, ɛ, ʌ, o, ɔ, ɑ/. In the later stages of L2 acquisition, SLM posits that even the L2 phones that fall in merged L1-L2 categories might be produced fairly well. Furthermore, L2 phones that are perceived as very dissimilar to an acoustically close L1 category that were produced poorly at early stages, with time, might be more accurately produced than L2 phones that are more similar to a close L1 phone. This is based on establishing new phonetic categories for the dissimilar vowels in the L2. PA learners of AE are supposed to perform in the same way for vowels such as /ʌ, ɔ, ɑ/, which are very dissimilar to any PA vowel.

According to Best & Tyler (2007), PAM-L2 specifically aims to describe and explain how the naïve nonnatives perceive the native speakers, while Flege's SLM and its revised version SLM-r focus on the production/perception of L2 sounds by L2 learners in the developed stages of the learning process. The SLM stresses the link between perception and production, suggesting that correct perception of phonetic distinctions between two L2 phones would eventually lead to correct production of these differences. These predictions were later validated with empirical evidence (e.g., Lengeris & Hazan, 2010; MacKay et al., 2006), attesting to the prediction of a relationship between perception and production that separates this model from other pure perception models, such as the PAM.

More recently, however, SLM-r focuses on perceived phonetic (dis)similarity among L1 and L2 phones or two L2 phones instead of "new" and "similar" phones (Flege & Bohn, 2021). Basically, this modification emphasizes the learner's role in perceiving the differences and similarities of the L2 contrast rather than revealing the acoustic differences between the L1 and L2 sounds. By focusing on the listeners' role, the perceived phonetic distance between segments, and learners' age of acquisition, the SLM-r will help explain the learners' differences within the context of the same L1 group acquiring the same L2 contrast, which applies to the case of this study on EFL. On the other hand, PAM is utilized to balance the framework using a direct perceptual scaling method as a viable option for categorizing listeners' perceptions.

One of SLM-r's properties is that it focuses on the coevolving between perception and production, as it hypothesizes that accurate perception of L2 sounds is no longer believed to precede accurate production (Flege & Bohn, 2021: 29) and that they should proceed in parallel.

However, as a pedagogical implication and since the two processes are based on the same cognitive resources, the focus at one time should be on one skill or the other and not both of them together.

Ultimately, I consider that a correct mental representation of what the phone should sound like is a necessary, but not sufficient, condition for correct/authentic pronunciation of the sound category because correct pronunciation does not automatically follow from a correct perceptual representation alone.

Both SLM and PAM models have substantially contributed to phonetics and phonology, especially in the phonological perception of nonnative sound segments. However, as with any other model, they have limitations. The previous version of SLM, although accredited by some practical studies (e.g., Bohn & Flege, 1990, 1992; Flege 1987), was criticized for not presenting clear criteria on how to differentiate between "similar" and "new" phones, especially in the case of vowels (Blankenship, 1991; Munro et al., 1996; Ingram & Park, 1997). SLM fails to make any explicit prediction concerning learners' ability to discriminate nonnative contrasts, which can be considered a major shortfall. It is also unclear how SLM should deal with prosodic (suprasegmental) differences between L1 and L2.

2.3.3. Second Language Linguistic Perception (L2LP) Model

Among the most recent models that address the perception and production of L2 is the Second Language Linguistic Perception (L2LP) model by Escudero (2005, 2009) and its revised version (Van Leussen & Escudero, 2015). The L2LP model consists of five theoretical ingredients (Escudero, 2005: 85-86), each of which covers a component of the L2 acquisition process theoretically and methodologically, viz. (i) the optimal perception, (ii) the initial state (first stage of L2 acquisition), (iii) the learning task, (iv) the development state, and (v) the end state (ultimate attainment). The five ingredients of the model provide explicit language learning predictions, linguistic explanations, and phonetic, phonological, and psycholinguistic constructs of L2 sound perception at the three logical (initial, development, and end) states of the acquisition process (van Leussen & Escudero, 2015).

In optimal perception, listeners maximize their efforts in understanding speakers by perceiving/categorizing L2 sounds through perceptual judgments that correspond to their intended message (Escudero 2005: 88), i.e., "an optimal listener will construct those vowels and consonants that are most likely to have been intended by the speaker" (Escudero 2009: 4). This mindset is applied in learning both L1 and L2 based on optimal perceptual mappings and sound categorization that extends to cover all the other stages of the L2LP theoretical ingredients.

At the initial state, it is predicted that learners will first perceive L2 sounds in a fashion resembling how they produce these same sounds in their L1 environment. In other words, the L2 initial state is the exact state of L1 at the time of starting to learn L2. The initial state is based on a comprehensive acoustic comparison between L1 and L2 that allows me to predict and explain L2 acquisition difficulties. L2LP assumes that learners are optimal perceivers of their L1 and that the L2 acquisition process begins with a full copy of the optimal L1 perception. At this point, predictions can be made on how the development of L2 will be shaped based on the acoustic similarities and differences between the two sound inventories. From this starting point, L2LP distinguishes between three scenarios to account for the predicted patterns of assimilation in *the initial state*, which are referred to as learning scenarios, i.e., the "NEW", "SIMILAR", and "SUBSET" scenarios:

- The L2LP NEW scenario mirrors the PAM SC assimilation and labels L2 contrasts that are acoustically close to typical or average tokens of a single L1 sound category. The L2LP NEW scenario is expected when two nonnative sounds are perceived as (and assimilated to) a single sound in the L1. For this scenario, the L2LP model expects a challenging learning problem because it predicts that learners will transfer their L1 perception grammar and sound categories (especially if L1 has fewer sound categories than are needed in L2) to fit the L2 environment, causing a nonnative realization of the affected sounds. Then, the scenario will involve the formation of new sound categories to connect the new categories to the perceptual integration of new acoustic dimensions, as it may involve integrating new auditory dimensions that have not been previously categorized.
- The L2LP SIMILAR scenario, unlike the previous one, applies when learners perceive two L2 phonemes as phonemically equivalent to (but phonetically different from) two L1 phonemes. The discrimination for this scenario is predicted to be easier with fewer learning problems than in the NEW scenario because in this scenario, i.e., SIMILAR, there will only be a mismatch between the L1 and the L2 perception of the two sounds in question.
- The third scenario is the SUBSET scenario, which occurs when one of the L2 phones is alternately perceived to belong to more than one category (phoneme) in L1. This is when the L2 category assimilation is undecided and falls between adjacent L1 sound categories because learners of L2 have a larger phonemic inventory in L1 than in L2.

As a bridge between the initial state and final state, the L2LP development stage proposes that L2 learning is accomplished in one of two ways: (i) learners will either create new categories for the perceived sounds, which will lead to new perceptual mappings, or (ii) they will modify the already existing perceptual mapping to fit both acquired languages. This mechanism underlies the learning process whereby learners move from the initial learning state to a final state. Therefore, L2LP hypothesizes that learners' mechanisms for L1 optimal perception will also underlie L2 linguistic perception.

In detail, each of the three L2LP scenarios comes with predicted solutions through learning tasks that should be performed to attain optimal L2 perception during the L2LP development stage. The NEW and SUBSET scenarios require two L2 learning tasks, i.e., a perceptual and representational task. However, the representational task for each scenario is suggested to be different. For example, if a contrast set is mapped to a single L1 category via an L2LP NEW scenario, then learners will need category-creation (or 'Category-Split') learning tasks that allow them to generate a new L2 category or split an existing L1 category to correctly perceive the new contrast. Based on the nature of this task involved in the NEW scenarios, it is predicted that the NEW scenario would be more difficult than the SUBSET scenario. A SIMILAR scenario will require learners only to reconsider already existing L1 categories and shift the L1 category boundary to fit within their optimal perception for L2 native location, which is predicted to be easier than the other two learning tasks.

To conclude, the L2LP learning task ingredient involves two subtasks, a perceptual task, and a representational task, to connect the three different states (initial state, the development, and the end state) of L2 sound acquisition. The perceptual task in a NEW scenario is based on shifting the already existing perceptual mapping of categories or on creating new ones for L2 mapping in comparison with those that have been duplicated from the L1 perceptual mapping. Meanwhile, a SIMILAR scenario involves only a boundary shifting process.

L2LP aims to completely describe, explain, and predict the L2 acquisition process at the sound perception level. The five stages of the model provide an understanding of each component of the L2 acquisition process. First, these ingredients initiate a comparison for the adult optimal perception across both languages involved via computations of the perceptual cue use. Second, based on this, the learning task will be set for L2 learners by fusing the LP model with its L2 extension. Finally, the third stage provides a proposal for the L2 development target, which is the predicted L2 learning task and its solution.

Since PAM is limited to explaining naïve nonnative and beginning L2 stages and while SLM targets the ultimate attainment of L2 speech learning (Tyler et al., 2014), L2LP differs from the previous two models in that it tackles the entire development process of L2 speech perception from naïve to native-like performance. Escudero (2005) based the L2 learning trajectories in L2LP on the optimal perception hypothesis, which claims that "the perceptual mapping of the signal depends on the particular characteristics of the listener's production environment" (Escudero, 2005: 52). Hence, the goal is to perceive L2 sounds by deciding to which category each sound belongs and minimizing the possibilities of misunderstanding a speaker (after Boersma, 1998: 337, 340, 371) based on maximum-likelihood behavior. Similar to PAM and SLM, L2LP predicts assimilation patterns, which are called "learning scenarios". However, contrary to the previous two models, these learning scenarios entail different learning tasks.

2.3.4. Predictions based on SLA theories and models

In a previous section (2.1.4), I provided a provision of different types of possible errors that PA learners of AE are hypothesized to commit based on the details discussed for each language. In this section, I discuss those predictions within the framework of each model that has been discussed in the previous sections. Table 2.10 provides a summary and a comparison of how each model and hypothesis categorizes speech sounds and predicts the relative difficulty of their acquisition.

Table 2.10. Relative difficulty of acquiring an L2 sound contrast as predicted by three SLA models; the cells in the table specify the scenario that is invoked at each level of difficulty.

Model	Relative degree of learning difficulty		
	Least	Moderate	Most
PAM	Two Category TC	Category Goodness CG	Single Category SC
SLM	Identical sounds	New sounds	Similar sounds
L2LP	SIMILAR	SUBSET	NEW

According to MDH, typological unmarked sounds that exist in the learner's L1 are supposed to be acquired effortlessly, while sounds that are more marked in the target language than in the source language are more difficult to learn. Therefore, if MDH is supported, most of the AE vowels that do not have a (near) counterpart in PA will be difficult to learn, since only the vowel qualities /i/ and /u/ are hypothesized to be unmarked and shared in both languages. If PAM is supported, vowels are to be divided into at least three categories, and the most difficult to acquire are vowels that are categorized with the SC scenario, followed by vowels with the CG scenario, and the least difficult vowels within the TC scenario. If SLM is supported by the results of this research, then the AE Similar vowels (e.g., AE /i, u/ with /iː, uː/ in Arabic) and consonants (e.g., AE /v, p/ and PA /f, b/) would be the most difficult to learn, and new sound (e.g., ŋ, ʤ) would be easy to acquire, while identical sounds would be the easiest (AE and PA

/t, d/). For the L2LP model, NEW sounds are considered the most difficult to acquire, while SIMILAR sounds are the least difficult, and the SUBSET case in L2LP might be applicable to PA consonants since Arabic has more consonants than AE, especially for the pharyngealized consonants (PA /$t^ʕ$, $s^ʕ$, $d^ʕ$/ confused with AE /t, s, d/, respectively). The prediction in such cases is of moderate difficulty.

CHAPTER THREE

PERCEPTUAL ASSIMILATION OF AMERICAN ENGLISH VOWELS BY PALESTINIAN-ARABIC LEARNERS OF ENGLISH

3.1. Introduction

3.1.1. Overview

This chapter investigates the perceptual assimilation of the 11 AE monophthongs through a PAM test by native Palestinian Arabic (PA) adolescents who learn English as a Foreign Language (EFL) in Palestine (experiment 1).

To the best of my knowledge, there are no previous studies that were conducted on the perceptual assimilation of AE vowels by PA learners or that have been devoted to an acoustic analysis of the difficulties that PA learners encounter in learning how to perceive AE vowels.

This study attempts to fill this gap by tackling PA learners' perception of AE vowels within the framework of the perceptual assimilation model (PAM & PAM-L2, Best, 1995; Best et al., 2001; Best & Tyler, 2007; Tyler, 2019). In light of the previously discussed models, the PAM framework is used to analyze the data and to classify all AE vowels (especially the problematic or difficult ones) in terms of categories of perceptual assimilation patterns.

3.1.2. Background of the study

This chapter of the thesis presents a study (study 1) on the relationship between the perceived phonetic similarity of AE monophthongal ('pure') vowels and the PA vowel inventory, to unravel how PA participants perceive the AE vowels as members of their L1 vowel inventory through a multiple-alternative forced-choice and category goodness rating task.

Many research articles have been written on the perception of L2 sounds. Most of the time, L2 adult learners do not perceive L2 sounds (especially vowels) the way native speakers of the target language do. The number of studies on L1 Arabic learners' perception and identification of English vowels is limited. Some studies have addressed different Arabic L1 speakers' perception, identification, and production of English vowels, especially of the major varieties, i.e., British English, American English, and Australian English (e.g., Almbark, 2012; Evans & Alshangiti, 2018; Faris et al., 2016; Nikolova-Simic, 2010), while some other studies were limited to sound production only by native Arabic learners (e.g., Ali, 2013; Hassan, 2014; Hubais & Pillai, 2010; Munro, 1993). These studies concentrated on bilingual Arabic-English speakers in general, and none of them approached PA speakers directly. The combination of American English vowels assimilated by Palestinian Arabic listeners in this study has not yet been covered in the literature so far.

Among the studies that dealt with the perception and production of English L2 vowels by L1 Arabic learners is Nikolova-Simic (2010), which focuses on the effects of dissimilarities between the Saudi Arabic and AE phonological systems and their impacts on learning vowels in EFL contexts by testing 20 beginners and 21 advanced-level Saudi Arabic learners of English L2. Her participants found it difficult to learn similar sounds that only differ in one feature, such as, /ɪ/ and /ʊ/ from their Arabic counterparts and more difficult to learn phonemic vowels in English that are considered to have allophonic counterparts in Arabic, such as /o/; the most difficult were those vowels that do not have a clear counterpart in Arabic, such as the mid vowels [ʌ], [ɛ], [e], and [ɔ]. More easily perceived and produced were the (almost) equivalent vowels between the two languages, i.e., /i/, /u/, and /æ/. The performance of the Saudi learners, both in perception and production, was not positively impacted by their experience, and there was a tight correlation between the perception and the production for most of the vowels. The results of the Nikolova-Simic (2010) study show that both beginners and advanced learners made errors mostly in the perception of vowels /ɛ/ (41% error) and /ɔ/ (43%). In vowel production, the two most mispronounced vowels were /o/ (56%) and /ɛ/ (48%). These results are in alignment with the results of a study on Iraqi Arabic, where the production of English vowels was not correlated to vowel perception as per Al-Abdely & Thai (2016a; 2016b), as discussed below.

In another study, Almbark (2012) investigated the perception and production of 18 Standard Southern British English (SSBE) vowels by 20 Syrian Arabic (SA) L1 learners based on the results of a previous pilot study (i.e., Almbark, 2011). The SA participants performed a perceptual assimilation task to match English vowels (produced by a native British English (BE) male speaker) with audible responses based on recordings produced by a Syrian Arabic male speaker. The results of Almbark’s (2012) PAM task show that SA assimilation patterns were distributed among TC contrasts (/æ~ɑː/, /ɔ~ʌ), SC contrast (/eɪ~ɛə/), CG (/ɪ~ɛ/, /əʊ~ɔː/), and a UC contrast (/iː~ɪə/). Generally, the PAM results in Almbark (2012: 372) evince that SA participants relied on the allophonic nature of vowels in their L1 in categorizing L2 SSBE vowels. The PAM task used in Almbarak (2012) was not the conventional way of performing a PAM test. In Almbarak (2012), the audible responses correspond to the phonemic vowels in

SA as well as to allophonic variants based on Cowell's (1964) discussion on the dialect of the city of Damascus. Moreover, these allophonic variations in SA are not guaranteed to be generalized to all SA participants. The learners may not be sensitive to the allophonic variations, especially naïve listeners, even if it was allowed to relisten to the responses a few times during the instruction phase of the experiment. Therefore, the results of written phonemic responses would not necessarily yield the same results as the audible allophones that Almbark (2012) presented as responses because participants' response**s** would be affected by their own renditions from their previous background if responses were presented in writing.

In another study, which included both perception and production of RP vowel sounds by Iraqi learners, Al-Abdely and Thai (2016b) tested 86 Iraqi L1 participants to discover the relationship between the development of perception and production of RP English vowels by L1 Iraqi Arabic learners. The study discusses different hypotheses on the relationship between perception and production of L2 phonemic contrasts while hinting at other factors that can weaken this link, such as L1 interference, proficiency level, age of learning, and length of residence. The same set of stimulus words used in Nikolova-Simic (2010) for Saudi Arabic L1 and Almbark (2012) for Syrian Arabic were reused in this study for the Iraqi learners. The participants underwent a placement test and were categorized into four groups according to their proficiency level following the Common European Framework of Reference for Languages (CEFR). The study consisted of two tasks, i.e., a perception task and a production task, and reported surprising results in the production task where learners from all groups had higher mean scores in the production task compared to the perception mean scores, except for one group, i.e., the A2 CEFR group, where participants showed significantly better results in the production of RP vowels than in perception, and the /ɪ, ʊ, ɔ, ɑː, oː/ vowels were produced significantly better than perceived, meanwhile, /æ/ was perceived better than produced. The results of this study claim a probable L1 interference because all the better produced RP vowels are found in the L1 variety vowel system, while the better perceived RP vowels do not have (near)equivalents in the L1 vowel system. Accordingly, the underlying assumptions postulated by models such as PAM and SLM that inaccurate perception results in inaccurate production of sounds were not well supported by Al-Abdely and Thai (2016b) based on the asymmetrical relationship between the perception and production results.

Since the mismatch between the Arabic and English vowel inventories is obvious, a one-to-one mapping between L1-Arabic and L2-English vowels is impossible. Therefore, I predict that either more than one L2 vowel will be assimilated to a certain L1 vowel, or some L2 vowels will go uncategorized. Moreover, these vowels could be uncategorized in a different manner.

In a study on vowel perceptual assimilation of L1-Arabic speakers of L2-English, Faris et al. (2016) focused on the different ways Egyptian Arabic (EA) learners assimilate the 19 Australian English (AusEn) vowels into the 10 EA vowels, including diphthongs and schwa based on Lehn & Slager (1959), while adding another 14 allophonic variants to “allow the EA listeners to be able to make choices reflecting fine-grained phonetic distinction” (2016: EL2). The results, first, show no systematic differences in participants’ assimilation of AusEn vowels to the EA phonemic categories and the allophonic categories (10 phonemic response alternatives + 14 allophonic response alternatives). All the additional allophonic variations that were used to yield fine-grained phonetic distinctions were reverted/collapsed into appropriate main phonemic categories of EA vowels (8 monophthongs and 2 diphthongs). Additionally, the results show that only 3 (out of 19) AusEn were assimilated into 3 EA vowels and categorized as follows: AusEn /ɐː/, /e/ and /o/ assimilated to EA /a:/ (87%), /i/ (60%), and /u/ (65%), respectively, and the rest of the AusEn were uncategorized. Counterintuitively, the EA /u:/ and /i:/ did not make it an assimilation category for any of the AusEn 19 vowels. EA /i:/ did not pass the 50% percent categorization threshold and only reached 49% but with a confident goodness rate = 5.64 (out of 7) for AusEn /i:/ which is the highest goodness rating in the whole study, but it was confused with AusEn /ɪə/ with 36% (4.86) and /ɪ/. Also, EA /u:/ was an uncategorized choice for 3 AusEn vowels: /ʉː/ with 35% (3.95), /o:/ with 21% (4.83), and /əʉ/ with 27% (3.91): not well categorized for any AusEn. The same situation applies for EA /a/ as it was assimilated to AusEn /ɐ/ (as in *calm*) with 45% (5.25) and to /æ/ with 30% (5.26). The results imply that EA participants confused their L1 vowel duration between long and short, as the results show no relation with AusEn tense and lax vowels, especially for uncategorized vowels. The results of this study reported different patterns within the uncategorized phonemes, which the authors relate to the perceptual relationship between L1 and L2 phonetic and phonological categories viz, focalized, clustered, or dispersed. Many AusEn vowels were reported to be uncategorizable for the native EA L1 learners, most of which were of clustered or dispersed types. Accordingly, the study strongly suggested that adult L2 English learners find it hard to assimilate (and therefore produce) accurate English vowels, in line with many studies in this domain (Best, 1994; Flege, 1995; Iverson et al., 2003).

Evans and Alshangiti (2018) focused on the perception of BE phonemes by 26 native Saudi Arabic learners living or working in London. Participants were divided into two categories based on their proficiency (high- and low-proficiency ‘HP’ and ‘LP’, respectively). They performed consonant and vowel identification tests, both in quiet and in noise. The study also included a production task to test whether the predicted difficulties in vowel perception affected participants' production as well and compared the results to native results. The study reports that some BE vowels with clear counterparts in Arabic were difficult to master even by experienced learners. For instance, both LP and HP participants confused BE lax /ɪ/ with BE tense /e/ (LP: 72%, HP: 44%), which have the same counterpart in Arabic, i.e., /i/. English high back vowels /uː, ʊ/ were found to be difficult to discriminate, although they have different (near)counterparts in Arabic. Meanwhile, vowels that do not have clear counterparts in Arabic were better discriminated, e.g., /e/ sound was correctly discriminated and was rarely confused with BE /ɪ/. In general, HP nonnative participants identified English vowels more accurately (69% correct) than the LP group (47%). Moreover, and as expected, the study demonstrated that the performance by both Arabic groups was poorer in comparison with native listeners and that difference was larger for vowels than for consonants. Nevertheless, poor performance was reported for both nonnative groups regardless of their proficiency level. Moreover, the study suggested a link between perception and production for the Saudi Arabic learners, as the results for participants with more accurate production were associated with better vowel perception.

3.1.3. Study questions

Considering these results of Arabic L1 studies on L2 English vowel (and consonant) perception and identification, it is evident that the relationship between Arabic and English is quite complicated, especially at the phonetic and phonological levels, so it is rather challenging to predict exactly what difficulties Arabic learners will have acquiring English as L2, especially when the Arabic low-variety is approached. Accordingly, one should recall the diglossic situation of Arabic as one of the main sources of these difficulties because there exists much variation between the low varieties of Arabic across the Middle East and North Africa (MENA) region, which will certainly affect the perception or assimilation relationship of a high-variety phoneme and its low-variety free variations. Thus, the difficulty of picking up nonnative phonemes may well be related to the L1 low-variety of Arabic of the speaker (different learning problems stem from different Arabic sound systems). Therefore, it is rational that the materials to be used in teaching EFL should be based on a comparison of each native language and the target language to spot the areas of difficulty for the said L1 language (Fries, 1945).

In light of the presented studies on English vowel assimilation and identification by different Arabic L1 dialects, I suggest the following:

1- Each L1 Arabic variety has its own characteristics that call for specific treatments to be adapted in teaching L2 pronunciation.

2- Different L1 Arabic varieties have different effects on L2 English vowel assimilation and identification, but some L2 English vowels are easily learned across all the mentioned L1 Arabic varieties.

As a whole, the overgeneralized results prompt the following specific questions for different Arabic varieties, in general, and for the Palestinian Arabic context in particular:

1- Which AE vowels constitute a (potential) perception problem for PA EFL learners?
2- (a) Which AE vowels are predicted to be the most difficult for them to perceive correctly?
   (b) How perceptually sensitive are PA learners to AE vowel duration and/or quality? and
3- How do the results of this research align with the results of similar studies on other CAVs learners of EFL?

From the studies discussed above, it is noted that their results suggest that some English vowels are grasped better than others and that not every nonnative sound is hard to grasp or creates a learning problem. Among the available theoretical models for such contexts, a cross-lingual model such as PAM can account for these variabilities in AE vowel perception (and production) across L1 Arabic varieties. As discussed earlier in Chapter 2 (section 2.3.1), this model focuses on the learner's degree of success in perceiving L2 AE vowels, which is subject to the perceived distance between the contrasting sound elements among PA and AE.

I have already highlighted the diglossic situation of Arabic nowadays and how MSA differs from CAVs even at the level of the phonological norm. Accordingly, it seems impossible to include all Arabic dialects within one study of English L2 vowel perception because predicting exactly how each dialect inventory responds is not straightforward or systematic across all dialects. In addition, different major English(es) have different vowel inventories in terms of quantity and quality. Consequently, this is an invitation for researchers to tackle other varieties of Arabic to discover how other Arabic learners' difficulties would vary according to the learner's local varieties and the specifics of their vowel inventory. Nevertheless, the neighboring varieties of PA (i.e., Levantine Arabic) may still share many phonological geo-sociological features and can benefit from this study.

3.1.4. Hypotheses

In light of a PA learner's environment, with little or no exposure to English and no direct interaction with native English speakers, PA learners' main source of learning English is formal education and, to a lesser degree, the (social) media. In addition, the PA vowel inventory – with all its allophonic capabilities – is still smaller than the AE vowel inventory, which will cause acquisition problems for the L1 PA participants. As discussed in the **literature review** chapter, many studies have argued that L1 learners rely on their L1 inventory when listening to an L2 speech and assimilate L2 sounds to the already available L1 categories (e.g., Best, 1995; Best, McRoberts & Goodell, 2001; Flege, 1995, 2002). Therefore, learners will find all AE vowels difficult to assimilate to their L1 inventory. Because of the mismatch between the two vowel inventories (in line with previous studies mentioned earlier on L1 Arabic vowels), no clean one-to-one mappings are possible.

More specifically, I hypothesize that the mid-high tense AE vowels /e, o/ will project onto AE /i/ and /u/, respectively, and will therefore be difficult to differentiate from the high tense vowels. Additionally, the AE lax vowels /ɪ/ and /ɛ/ may assimilate to PA short /i/. Also, this may be the case for AE /ʊ/ and /ʌ/ assimilating on PA /u/. A vast confusion is predicted for all (mid)low vowels, whereby /æ, ɔ, ɑ/ will all project onto long PA /a:/. Alternatively, /ʌ/ and /ɛ/ may project onto short PA /a/, in which case AE /ʌ/ and /ʊ/ would form a TC contrast (see Figure 2.3 in subsection 2.1.4.1), which should be easy to acquire.

Moreover, the durational cue of short/long vowels is a contrastive feature in the Arabic vowel inventory, i.e., duration is more salient acoustically than any spectral differences (quality). Such that PA learners will not be sensitive to vowel quality if the vowels are not spectrally remote from each other, such as will be the case between pairs of adjacent vowels in American English, especially in the midsection. Therefore, I hypothesize that PA learners will most likely depend on the temporal cue more than the quality of a vowel to discriminate between lax and tense vowels, while English speakers would primarily attend to the quality difference (Bohn, 1995; Flege et al., 1997; Hillenbrand et al., 2000; Van Heuven et al., 2020). Accordingly, the assimilation of AE vowels by PA learners will be affected by this process and probably would yield a less accurate perception than what will be observed for native English users.

3.2. Method

3.2.1. Preliminaries

An AE vowel assimilation test (as detailed in section 3.2.4) was administered to the native PA group to determine how they categorize the 11 AE monophthongs in terms of their own six L1 Arabic vowel categories. Before the data collection started, approval to conduct all the experiments of this dissertation was obtained from the Ministry of Higher Education and Scientific Research in Palestine. Additionally, I received a confirmation from the Ethical

Committee of Pannonia University to conduct the experiments under the university affiliation. All experiments on PA took place in two secondary schools (As-Salahiya Boys' Secondary School and Al-Aishiya Girls' Secondary School) in Nablus city to the north of the West Bank, Palestine.[22]

3.2.2. Stimulus material

The stimulus recordings were selected from Wang & Van Heuven (2006, 2014) for two native AE male speakers from a larger group of 20 participants. The two selected speakers were the only two who properly differentiated between all pairs of American English monophthongs, including the (mid-)low back vowels found in only some AE dialects (Carley & Mees, 2019; Ladefoged & Disner, 2012; Strange et al., 2007), i.e., /ɑ/ and /ɔ/. Appendix 3.1 shows an acoustic analysis of the monophthongs produced by the two speakers. The two speakers were recorded on digital audio tape (DAT) in a sound-insulated recording booth through a Sennheiser MKH-416 microphone. The materials were later downsampled (16 kHz, 16 bits) and stored on a computer disk. For each speaker, the set of 11 monosyllabic words or phrases produced in /hVd/ carrier word, was excerpted from the fixed carrier *Now say…… again* (yielding 22 tokens) via the digital waveform editor in the Praat speech processing software (Boersma & Weenink, 2021; Boersma & Van Heuven, 2001). The carrier words are: heed /i/, hid /ɪ/, hayed /e/, head /ɛ/, had /æ/, hud /ʌ/, hod /ɑ/, hawed /ɔ/, hoed /o/, hood /ʊ/, and who'd /u/, following the established practice in, e.g., Peterson and Barney (1952) and Hillenbrand et al. (1995).

3.2.3. Participants

Prior to testing, the participants completed a language background and attitudes questionnaire (LBQ) (after Flege et al.,1995; 1997; 1999). The PA participants (throughout the dissertation) were 40 native PA high-school students (20 males, 20 females). Their ages ranged from 15;10 to 17;1 years old (*Mean* = 16;4) at the time of conducting the experiments (see Appendix 3.2). All participants had learned English as a foreign language (EFL) as part of their formal education since first grade and were eleventh graders (junior year) at the time of conducting the experiments presented in this dissertation. All the PA participants were living in Palestine when they performed the experiments. Their native language is Arabic, their parent's native language is Arabic, their birthplace is in an Arabic-speaking region, they had never lived in a non-Arabic

[22] I would like to thank all the participants, teachers, and school principals who facilitated the progress of my experiments. I would like also to thank Ministry of Education and higher Education in Palestine, represented by its minister Prof. Dr. Marwan Awartani, for their help in approving and facilitating the experiments.

speaking region for a period of over 3 months, and the only language they had ever been immersed in is Arabic. The PA participants reported using Arabic more than English at home (4.5 vs. 4.0 on a 7-point Likert scale), at school (4.4 vs. 4.0), with friends and at parties (4.8 vs. 4.2) and for the last five years before conducting the experiments (4.6 vs. 3.8).

The level of education of their parents varied as follows: 50% have a Bachelor's degree or higher, 42% finished secondary school, 5% finished primary school, and 3% finished elementary school. The socioeconomic situation of the parents varied between 22.25% government employees, 12.5% private sector employees, 21.25% self-employed, 2.5% daily workers, and 41.5% unemployed or housewives. The participants' legal guardians signed an informed consent form (see Appendices 3.3.A-B) on behalf of the participants, as the participants themselves were under the age of legal capacity, the participants themselves approved their participation orally, and they were informed that they could quit the experiment at any time before or during the test. All the participants reported having and demonstrated normal speech and hearing without any problems and considered Arabic their L1. The participants participated voluntarily and were not remunerated. To maintain the sample homogeneity and preserve the study limitations (and credibility), three factors defined the inclusion criteria: participants were included if they had not had any English speaking or listening training courses, considered Arabic their L1, and had not lived in an English-speaking country for more than 3 months.[23] Participants completed the questionnaire within approximately 10 minutes.

We have no expert evaluation of the quality of the student's individual pronunciation. This aspect is not formally assessed in the Palestinian secondary school system (see Farran et al., 2020; see also Introduction). The students' marks (i.e., teacher judgment) are mainly based on quality of written work and/or reading comprehension tests. The curriculum requirements stipulate that, by the end of the last year, the student should speak intelligibly such that a native English listener will understand the student with little effort. More than likely, the student will speak English with a strong Arabic accent, substituting Arabic sounds for the English targets with hardly any modification. Consonant clusters will be modified by deletion and epenthesis to fit the more basic Arabic syllable structure. Similarly, the stress patterns will be wrong (inspired by Arabic stress rules), and all content words will have sentence stresses. In the Language Background Questionnaire (LBQ), the students were asked to what extent they agreed (1 'fully agree' to 7 'fully disagree') with six statements (at different points in the

[23] Two potential (male) participants were excluded from the whole study as they reported to have lived in the USA for a period longer than 3 months.

questionnaire) that pertained to their pronunciation of English. The six statements are listed in the table below, with the means and standard deviations in the responses. For the sake of comparison, I added the responses to two similar questions about their pronunciation of Arabic. Table 3.1 provides a summary of the participants' response to the mentioned questions. The questionnaire also (see Appendices 3.4.A-B) obtained information about the participants' self-evaluation of specific language abilities (e.g., grammar, communication, pronunciation), attitudes toward their acquired languages, how often they used them, and in what contexts they were exposed to each language. The LBQ also required participants to self-evaluate their proficiency in Arabic and English.

Table 3.1. Responses by 40 adolescent Palestinian Arabic learners of English as a foreign language to 8 statements. Learners expressed their agreement/disagreement with each statement on a 7-point scale with 1 = 'fully agree' to 7 = 'fully disagree'. Means and standard deviations are specified overall and for boys and girls separately. The effect of gender is evaluated by t-tests for independent samples (df = 38 in all tests, two-tailed testing).

Statement	Gender						t-test	
	Male		Female		Total		t	p
	Mn	SD	Mn	SD	Mn	SD		
1. I pronounce English well	4.45	2.16	2.00	1.30	3.23	2 15	4.3	<.001
2. People praise my ability to pronounce English	4.05	2.09	2.70	1.69	3.38	2.00	2.3	.031
3. I am good at imitating foreign accents and dialects	4.55	1.85	2.50	1.64	3.53	2.01	3.7	.001
4. I have a good memory for the way words are pronounced	4.30	2.32	2.25	1.71	3.28	2.26	3.2	.003
5. It is important to pronounce English well	5.00	2.64	1.55	1.28	3.28	2.69	5.3	<.001
6. Pronunciation doesn't affect how well you can communicate	3.95	2.52	4.50	2.37	4.23	2.43	–.7	.481
7. I pronounce Arabic well	4.70	2.74	1.65	1.23	3.18	1.60	4.6	<.001
8. It is important to pronounce Arabic well	5.20	2.24	1.70	1.13	3.45	2.49	6.2	<.001

On the average, the students are non-committal with a mean response close to the midpoint of the scale (which is 4). However, the effect of gender is large. The girls agree much more with the statements (except for #7) than the boys, which indicates that the girls consider their pronunciation of English, as well as their talent and motivation for learning EFL pronunciation, better (and in the upper half of the scales, i.e., between 1 and 4) than the boys (who rate their abilities, talent and motivation in the lower part of the scales (i.e., between 4 and 7). Item #7 was formulated as a negation, and probes whether the student thinks that a good EFL pronunciation is irrelevant for successful communication. Here the girls tend to disagree more strongly with this negative view than the boys, but the difference falls short of significance.

Somewhat disconcertingly, the same difference between the boys and girls, but in an even more extreme form, can be observed with respect to their L1 Arabic. The boys believe they pronounce Arabic rather poorly, and do not consider a good pronunciation of Arabic important, while the women rate their native pronunciation positively and attach great importance to this. It is well documented in sociolinguistic studies that women attach greater importance to correct pronunciation of the native language than men in so far as they are conscious of pronunciation rules for the standard language (Labov, 2001). Greater motivation on the part of women to acquire a good pronunciation in foreign languages has often been reported as well, but the discrepancy between the genders in the attitudes towards good pronunciation both of L1 and L2 in the present group of Palestinian Arabic EFL learners seems unprecedented.

3.2.4. Procedure

The task was explained in the listeners' native PA language by a PA native speaker (the author). The stimulus recordings were presented to participants individually as they sat at a desk in a small-sized quiet room at a comfortable listening level via a good quality headset (Sennheiser PC150) through a noiseless notebook computer and a Praat MFC script. The Praat script controlling the instructions, stimulus presentation and data collection is provided in Appendix 3.5 for this task. Figure 3.1 shows snapshots of the test interface that the PA participants saw in the different phases of the experiment.

Figure 3.1 Screenshots showing the different stages of the PAM test in integrated Arabic text. Top-left: intro. Top-right: six response categories activated [in yellow]. Bottom-left: 5-point goodness rating scale activated after choosing a category. Bottom-right: test-end slide and a *thank you!* message.

The task was self-paced with a closed-set format and was presented in full-screen mode. In this task, 44 AE stimuli (11 monophthongs for the two native AE speakers, repeated twice) were made audible only once and one at a time. The response buttons contain meaningful (real) words written in Arabic script in a fixed /sV(:)r/ carrier which only differ in the nucleus vowel to cover all Arabic short/long vowel inventory. The Arabic responses read: /sa:r/ 'walked', /sar/ 'to rejoice', /su:r/ 'outside-/outer wall', /sur/ 'rejoiced', /si:r/ 'walk', and /sir/ 'secret'.

Accordingly, each button represents one of the six Arabic monophthongs. The Arabic words were integrated within the Praat MFC script using Arabic symbols but in reversed order in the script so that in the interface they appear in correct order, and without the need to present the words in images above the buttons. This integration allows the participant to see how her/his choice is highlighted (while other choices are dimmed) and shift their attention to the goodness rating scale that became activated after choosing a category. The response buttons were arranged in two columns reflecting the length of the vowels (left: short vowels, right: long vowels) and three rows corresponding to vowel quality (top: back vowels; middle: mid vowels; bottom: front vowels).

Before the actual test, participants listened to the task instruction in Arabic while the instructor showed them a PowerPoint presentation explaining each step of the task. In an initial familiarization phase. Participants tried the task a few times to get used to it; afterwards, they performed the actual task. The actual task took approximately 20–25 min to complete. The instructions for this task were to listen to each played stimulus and decide as quickly as possible which of the L1 vowels the heard token is closest to. After choosing the most representative/appealing vowel (represented in a word), the chosen word was highlighted (changed color from yellow to red). At the same moment, the goodness scale was activated (from gray to yellow), prompting the participant to rate the token on a goodness scale from 1 to 5 (5 = good). After hearing the stimuli, choosing a word, and rating the token, the interface was reset, and the next token was made audible after 2 seconds. The process was repeated until all 44 stimuli were responded to. Stimuli were presented in a quasi-random order, excluding immediate repetition of the same token. The order was different for every participant. The response latency was measured (in milliseconds) from the onset of the auditory token until the goodness rating button was clicked. Reaction time was excluded from the analysis in this thesis. Participants were able to monitor their progress following the counter at the top left of the interface. A *thank you* message was shown upon finishing the task. No other feedback was given. The task was administered through a Praat MFC script that presented stimuli and collected responses (choice, goodness rating, and response latency). The total number of responses at the end of this task was 1,760 (40 PA participants × 44 stimuli). Statistically speaking, there are no hypotheses in this test which merely collect data, from which learning problems can be predicted. Therefore, what could be considered dependent variables in this test are: Arabic vowel category, goodness judgment, and response latency (which was excluded from the analysis). Also, only one variable is considered an independent variable: the AE vowel category (11 types).

3.3. Results

Table 3.2 shows the perceived phonetic similarity between 11 AE vowels (rows) and 6 Arabic vowels (columns) by PA participants. Given that the 40 participants heard each token 4 times, the maximum number of votes for each AE vowel is 160.

Table 3.2. Perceptual Assimilation Model test results.

Stimulus		Statistics		Response												Total
				i:		i		a:		a		u		u:		
heed	i	N	Typ.	**153**	**(4.3)**	7	(3.9)									160
		%	Fit	**95.6**	**4.1**	4.4	.2									100
hid	ɪ	N	Typ.	15	(3.7)	**144**	**(3.9)**	1	(4.0)							160
		%	Fit	9.4	.0	**90.0**	**3.5**	.6	.0							100
hayed	e	N	Typ.	**145**	**(4.0)**	11	(3.7)	4	(3.0)							160
		%	Fit	**90.6**	**3.6**	6.9	.0	2.5	.1							100
head	ɛ	N	Typ.	10	(3.8)	**142**	**(4.2)**	3	(3.0)	5	(3.6)					160
		%	Fit	6.3	.0	**88.8**	**3.7**	1.9	.0	3.1	.1					100
had	æ	N	Typ.	9	(3.7)	16	(3.6)	**131**	**(4.4)**	4	(4.0)					160
		%	Fit	5.6	.2	10.0	.4	**81.9**	**3.6**	2.5	.1					100
hawed	ɔ	N	Typ.			1	(4.0)	18	(4.5)	2	(2.5)	13	(3.5)	**126**	**(3.9)**	160
		%	Fit			.6	.0	11.3	.5	1.3	.0	8.1	.3	**78.8**	**3.1**	100
hod	ɑ	N	Typ.			1	(4.0)	**98**	**(3.6)**	6	(3.2)	5	(3.6)	50	(3.2)	160
		%	Fit			.6	.0	**61.3**	**2.2**	3.8	.1	3.1	.1	31.3	1.0	100
hud	ʌ	N	Typ.			5	(3.2)	1	(2.0)	8	(3.8)	**140**	**(4.3)**	6	(3.8)	160
		%	Fit			3.1	.1	.6	.0	5.0	.2	**87.5**	**3.8**	3.8	.1	100
hoed	o	N	Typ.	1	(4.0)			3	(3.7)	1	(3.0)	7	(3.4)	**148**	**(4.3)**	160
		%	Fit	.6	.0			1.9	.1	.6	.0	4.4	.1	**92.5**	**4.0**	100
hood	ʊ	N	Typ.	1	(5.0)	6	(3.0)			1	(5.0)	**139**	**(4.1)**	13	(3.9)	160
		%	Fit	.6	.0	3.8	.1			.6	.0	**86.9**	**3.6**	8.1	.3	100
who'd	u	N	Typ.	1	(3.0)			1	(2.0)	8	(3.8)	3	(3.7)	**156**	**(4.4)**	160
		%	Fit	.6	.0			.6	.0	5.0	.2	1.9	.1	**97.5**	**4.3**	100
Total		N		335		333		259		27		307		499		1760
		%		19.0		18.9		14.7		1.5		17.4		28.4		100

Key: Absolute (N) and relative (%) number of responses, typicality rating (Typ. on a scale from 1 to 5 = best, in parentheses), and fit index (= (Typ. × %)/100) of the categorization of American-English vowels into Palestinian Arabic vowels. Shaded cells in green and boldface contain the preferred ('modal') response category. Fit index in green cell: Good, in yellow cell: Fair, in red cell: Poor, in gray cell: Unclassified (for details see text).

Each cell of Table 3.2 lists four statistics: N (absolute number of responses), Typ. (mean typicality rating), % (relative number of responses, expressed as the row percentage), Fit (Fit index = ((Typ. × %)/100). Each category percentage expresses how attractive the perceptual assimilation is to a certain AE vowel category for the average PA listener, while the mean rating provides a measurement of the perceived goodness of this AE vowel as a token of the PA category. Cells highlighted in green and boldfaced contain responses with 50% or better agreement among the listeners. When the number of responses in a cell is between 25 and 50%, the cell is highlighted in yellow. This occurred only once. Cells without shading contain less than 25% of the responses. If a response category was never used by the participants, it was left empty in the table.

As a general rule, I consider any nonnative vowel Categorized (C) when it is assimilated to a native category with ≥ 50% of the responses. If two (or more) response categories receive a total above 50% of the responses and each category has at least 25% of the responses, then I consider the nonnative vowel Uncategorized (U). In all other cases, the nonnative vowel is considered Non-assimilable (N).

Accordingly, each of the eleven American English vowels was found Categorizable in terms of the (Palestinian) Arabic vowel inventory. There is a green cell in each row in Table 3.2, which means that each AE monophthong maps onto one of the Arabic vowels with an agreement better than 50%. However, PA short /a/ did not qualify as an assimilation category for any of the 11 AE vowels, i.e., no English vowel comes close enough to it. This can be justifiable on the basis that the two low AE vowels are tense and therefore long. That is why the /a:/ alternative is preferred by the participants in the case of AE low vowels. Meanwhile, all other vowels are assimilated in pairs, either in an SC scenario or in a CG scenario with high assimilation percentages (> 60%, percentages mean = 86.49%).

The statistics in Table 3.2 show how 10 of the 11 AE monophthongs were assimilated to a single Arabic vowel with relatively high percentages (79–98%) and with good typicality ratings (3.9–4.4). The only exception is AE /ɑ/ (hod), which was identified as PA /a:/ in 61% and as PA /u:/ in 31% of the responses, with typicality ratings of 3.6 and 3.2, respectively.

In this section, I consider all pairs of AE vowels and classify them into scenarios. Given the inventory of 11 AE monophthongs, there are 55 possible contrasts that EFL learners should be able to make. The large majority of the contrasts in Table 3.3 (i.e., 47 of the contrasts) are of the TC type (in light green cells) and should present no problem to EFL learners with a Palestinian-Arabic L1 background. They are unproblematic because each member involved in the contrast maps onto a clearly different vowel phoneme in the learner's native language. For instance, the contrast between *heed* and *hid* is of the easy TC type because *heed* maps onto the long/tense vowel /i:/ in Arabic while *hid* maps onto the short/lax /i/ vowel.

Table 3.3. Contrast matrix showing the 55 pairwise contrasts for American English vowels assimilated to Palestinian Arabic in terms of PAM categories. The contrast type is indicated in the cells (TC: Two Categories, SC: Single Category, CG: Category Goodness).

			i	ɪ	e	ɛ	æ	ʌ	ɑ	ɔ	o	ʊ
			C	C	C	C	C	C	C	C	C	C
heed	i	C i: 96										
hid	ɪ	C i 90	TC									
hayed	e	C i: 91	**SC**	TC								
head	ɛ	C i 89	TC	**SC**	TC							
had	æ	C a: 82	TC	TC	TC	TC						
hud	ʌ	C u 88	TC	TC	TC	TC	TC					
hod	ɑ	C a: 62	TC	TC	TC	TC	**CG**	TC				
hawed	ɔ	C u: 79	TC	TC	TC	TC	TC	TC	**SC**			
hoed	o	C u: 93	TC	TC	TC	TC	TC	TC	TC	**SC**		
hood	ʊ	C u 87	TC	TC	TC	TC	TC	**SC**	TC	TC	TC	
who’d	u	C u: 98	TC	TC	TC	TC	TC	TC	TC	**SC**	**SC**	TC

However, seven contrasts are predicted to be difficult to learn, as shown in Tables 3.1–3. These are the SC (Single Category) pairs indicated in the red cells in Table 3.3 (green cells of the same column in Table 3.2). For instance, vowels in the SC contrast *heed* ~ *hayed* will be difficult to discriminate for Arabic EFL learners because both are heard as convincing exemplars of the same Arabic vowel category, i.e., long/tense /i:/. There are another six difficult-to-learn SC contrasts. The AE (hid /ɪ/), and (head /ɛ/) (lax, front-high, and mid-low vowels in English) assimilate equally well to Arabic high, front, and short /i/. Accordingly, this contrast predicts great learning problems because of the difficulty in discriminating between the two AE vowels. The same assimilation SC scenario is found for the AE pairs (hayed /e/) and (heed /i/) as both assimilate to the Arabic high, front, and long /i:/; (hood /ʊ/) and (hud/ʌ/) both assimilate to Arabic high back and short /u/. All the SC contrasts (see Table 3.4) are predicted to be difficult to discriminate by the PA learners and will require more time and effort to correctly subdivide the L1 category into two new and smaller categories that correspond to the L2 categories. However, learners are predicted to adjust their perception of these contrasts after the development of their L2 faculty and either form new phonetic categories for one member of each problematic set of AE vowels before being able to create new phonological contrasting categories for both of them or subdivide the already existing L1 category into two categories that correspond to the L2 vowels (Best & Tyler 2007).

Table 3.4. Summary of problematic contrasts based on PAM test results.

	AE/L2 contrasts	AE vowel	PA/L1 equivalent	PAM assimilation type
1.	had	æ	/a:/	CG
	hod	ɑ		
2.	heed	i	/i:/	SC
	hayed	e		
3.	hid	ɪ	/i/	SC
	head	ɛ		
4.	hood	ʊ	/u/	SC
	hud	ʌ		
5.	who'd	u	/u:/	SC
	hoed	o		
6.	who'd	u	/u:/	SC
	hawed	ɔ		
7.	hoed	o	/u:/	SC
	hawed	ɔ		
8.	hawed	ɔ	/u:/	SC
	hod	ɑ		

Finally, there is one contrast that answers to the CG (Category Goodness) scenario according to Tables 5.1–3. This is the contrast between the two low vowels in *had* /æ/ and *hod* /ɑ/ (front and back, respectively). They both map onto front, low long /a:/ in PA but the match is clearly better for *had* (82% match) than for *hod* (61% match). In fact, *hod* maps onto Arabic /a:/ (61%) as well as onto Arabic /u:/ (31%). This makes sense, of course, since AE /ɑ/ is in between PA /a:/ and /u:/, but closer to the former than to the latter, which explains the better match with PA /a:/. The prediction of learning difficulties for contrasts in the CG scenario is that the learners will soon grasp the difference between the contrasting set and discriminate between them (however, not as well as they do for the TC type of assimilation) because they will be sensitive to the difference in typicality between the two sounds in the mapping onto the same category in the L1; they will be aware of the fact that the vowel in *had* is closer to their native /a:/ than the vowel in *hod*. Among predictions for such a type of assimilation is that learners will form a new phonetic and phonological category for the deviant AE vowel with a lower fit index rate, i.e., *hod* /ɑ/ different from PA /a:/, while the better AE exemplar of the PA vowel, i.e., /æ/, will remain phonetically and phonologically perceived as a good representative of PA /a:/ without forming any new category for it. From another perspective, this process can be seen as a splitting movement where learners split their L1 /a:/ category into two different categories reflecting the differences between AE /æ/ and /ɑ/.

A few things should be noted based on the results here. First, in this test, the vowels assimilated to PA were never confused in terms of vowel length. Both members of the problematic assimilation pairs (SC or CG) were assimilated to either short Arabic vowels or long Arabic vowels. Figure 3.2 below shows how the PA assimilation contrasts are distributed over the AE IPA vowel chart. There is a clear stratification as the AE tense vowels assimilated only to PA long vowels and the AE lax vowels assimilated only to short PA vowels. Even for these confused/problematic contrast sets, the confusion is not in duration but between spectrally adjacent vowels (i.e., differing in vowel quality).

Second, it should be highlighted that the identical phonetic symbols used in this chart do not mean they have identical phonetic qualities cross-linguistically between AE and PA. Therefore, the vowel space that PA /i/ sound occupies, for example, is different, probably much larger, from (and was not used as a phonological identical equivalent to) the vowel space reserved for AE /i/. Their renditions are very different between the two languages.[24]

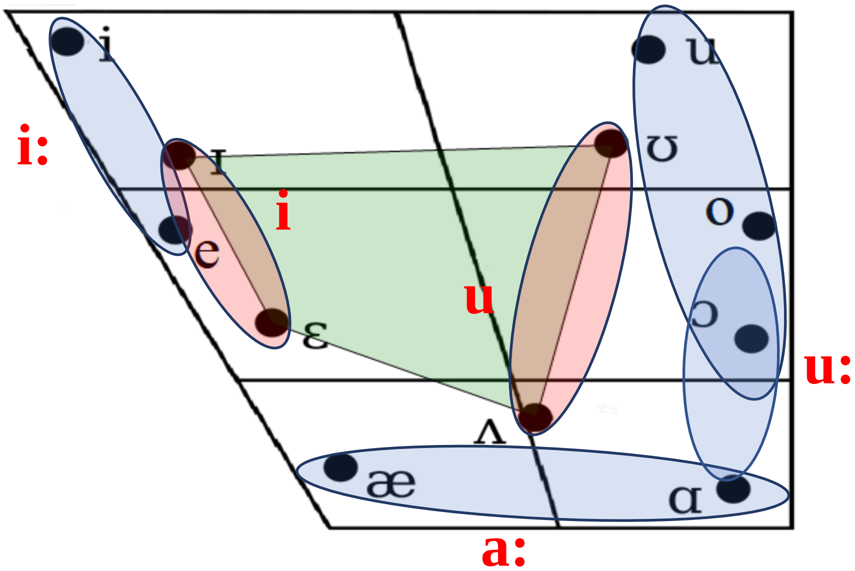

Figure 3.2. Pairs or triplets of AE monophthongs that map onto the same vowel in PA. Blue ellipses: long/tense vowels, red ellipses: short/lax vowels.

As discussed earlier, learners did not perceive any of the AE monophthongs as a response category for short /a/, however, the other PA vowels are assimilated to in an organized way. The (potentially) problematic AE vowel pairs contain spectrally adjacent vowels that are almost equivalent to one of the PA vowels, which justifies the fluctuations in the goodness ratings.

Third, the typicality judgments are only weakly correlated with the agreement obtained on a particular response category, r = .398, N = 44, p = .004, one-tailed). Although significant, the correlation accounts for only 16% of the variance. This means that the percentage of votes for

[24] Note that in British English (or other varieties of English in which ash is a short vowel, the contrast between *had* and *hard* (but not *hod*) would be immediately clear. In Standard Southern British English (SSBE) the /r/ in *hard* is not pronounced, and there is a contrast between long open /ɑː/ and short half-open /ɔ/.

a particular response category and the typicality rating provide largely independent information on the likelihood that an L2 sound will be accepted as a token of some native sound category.

Another way PAM considers the perceptual assimilation of L2 sounds to L1 is by categorizing individual nonnative sounds instead of categorizing them into scenarios. To do justice to the relative independence of agreement and typicality in PAM-L2, some studies (Guion et al., 2000; Best & Tyler, 2007) suggest that the information provided by agreement and typicality rating be combined into a single measure, which they call the fit index (a predictor of how well a foreign sound fits a native category) that provides "categorial discrimination of the L1-L2 phone pair using multiple speakers of each language" (Best & Tyler, 2007: 28). The fit index is computed for each cell in Table 3.2 by multiplying the goodness rating by the relative number of votes the response category received (i.e., by the row percentage).

Table 3.5 copies the fit indexes for all 11 AE monophthongs as calculated for each L1 classified vowel that received more than 25% of the relative number of responses (see % number in Table 3.2). The combination of the identification percentages and the goodness rating provides one single metric that weighs the proportion of the classification of each member of the contrasting sets to its PA classified category.

Table 3.5. Fit indexes and PAM categorization of the AE problematic pairs for PA listeners, in descending order of fit index.

AE vowel	PA vowel	Mean goodness	Modal Choice (%)	Fit index	Decision
u	/u:/	4.4	97.5	4.3	Good PA /u:/
i	/i:/	4.3	95.6	4.1	Fair PA /i:/
o	/u:/	4.3	92.5	4.0	Fair PA /u:/
ʌ	/u/	4.3	87.5	3.8	Fair PA /u/
ɛ	/i/	4.2	88.8	3.7	Fair PA /i/
æ	/a:/	4.4	81.9	3.6	Fair PA /a:/
ʊ	/u/	4.1	86.9	3.6	Poor PA /u/
e	/i:/	4.0	90.6	3.6	Poor PA /i:/
ɪ	/i/	3.9	90.0	3.5	Poor PA /i/
ɔ	/u:/	3.9	78.8	3.1	Poor PA /u:/
ɑ	/a:/	3.6	61.3	2.2	Unclassified
ɑ	/u:/	3.2	31.3	1.0	Unclassified

The fit index was calculated following Guion et al. (2000). As a working hypothesis (Wang & Chen, 2019, 2020), the fit index was categorized into three brackets (Good, Fair, and Poor exemplar of the L1 vowel) using the mean of the modal choices and the standard deviation as a criterion. The mean fit index obtained for the 5 PA vowels = 3.8 (SD =.9). Accordingly, I consider each AE vowel with a fit index ≥ 1 SD above the mean (i.e., ≥ 4.3) a "Good" instance of the PA vowel chosen for its category. Any AE vowel with a fit index between the mean and

+1 SD (i.e., between 3.8 and 4.3) is considered a “Fair” exemplar of the PA vowel category. AE vowels between the mean and 1 SD below the mean fit index (i.e., between 2.9 and 3.8) are classified as “Poor”. AE vowels with a fit index less than 1 SD below the mean are considered unclassified.

Within the problematic assimilation sets, Table 3.5 shows which member of each contrast pair is sooner perceived as an exemplar of the PA vowel. In the PAM context, vowels can be categorized as either a good, fair, or poor exemplar, uncategorized, or non-assimilable (Tyler, 2019). Then, categorized pairs can be found to fall in either a TC, CG, or SC scenario. The results in Table 3.5 show which member of each pair is more similar to the L1 PA vowel, especially for contrast pairs with close voting/responses and goodness rating. The opposite is also true because it tells how remote an AE vowel is from a PA vowel it assimilated to compared to the other member of the contrast pair.

Let us now take a closer look at the potentially problematic assimilation categories. Table 3.5 shows how all the contrast sets are weighted with their fit indexes. The CG set for the AE /æ/ and /ɑ/ vowels shows that they both assimilate to PA /a:/ but with varying fit indexes. AE /æ/ is perceived as a better (and *Good*) exemplar of PA /a:/ with a 3.6 fit index than AE /ɑ/ with a 2.2 fit index, which demotes it to *Poor*. The members of the SC pair /i/ ~ /e/ (*heed* ~ *hid*) both assimilate to (and both were rated as Good exemplars of) PA /i/, but with a bias towards AE /i/ with a fit index of 4.1 compared to 3.6 for AE /e/, i.e., /i/ (but not /e/) is a better exemplar of PA /i:/ even if they both were found to be in an SC contrast. Likewise, both AE /ɛ/ and /ɪ/ are assimilated to PA /i/ with almost equal fit indexes (3.7 and 3.5, respectively), yielding another SC scenario. The same is observed for AE /ʌ/ and /ʊ/ as both are perceived as *Good* instances of PA /u/. Accordingly, all these SC scenarios are rated with *Good* fit ratings of their PA assimilated to vowels; each pair assimilates to a single PA vowel. The exception is the /ɔ/~ /ɑ/ contrast. Although it was found to be within an SC contrast to PA /u:/, the fit index for AE /ɑ/ is the lowest exemplar of PA /u:/ and of any PA vowel.

In contradistinction to this, PA /u:/ was the assimilation choice for four AE vowels with different SC contrasting pairs. AE /u/, /o/, /ɔ/, and /ɑ/ were assimilated to PA /u:/ in a descending order ranging in fit index rating between *Good*, *Fair*, and *Poor*, respectively.

The *Good* ranking of AE /u/ with a 4.3 fit index and AE /o/ with a 4.0 suggests that they share some properties with PA /u:/, more than the other assimilated AE vowels in this context but with a better fin index to /u/ than /o/. Actually, AE /u/ was the highest AE vowel with a *Good* categorization as a PA vowel in general and as PA /u:/ in particular. To a lesser degree, AE /ɔ/ was still a *Fair* exemplar of PA /u:/, better than /ɑ/.

These results are in line with previous studies, especially with respect to AE /ɑ/ and /ɔ/ (Al-Abdely & Thai, 2016a; Faris et al. 2016; Nikolova-Simic, 2010). Different studies on the perception of English vowels by Arabic dialect speakers have reported that AE English /ɑ/ (/ɒ/ in SSBE and RP) was the most difficult vowel to identify correctly (Al-Abdely & Thai, 2016a: 9; Almbarak, 2012: 267; Evans & Alshangiti, 2018: 21, 25; Nikolova-Simic, 2010). Even though this study shows that AE /ɑ/ and /ɔ/ are among the most difficult vowels for PA learners, the assimilation pattern of AE /ɑ/ is different from what was found in other L1 Arabic studies. For instance, in Almbarak (2012), /ɑ/ (or SSBE /ɒ/) assimilates to two Syrian Arabic short vowels (/o/ and /a/) with better assimilation to short /o/ than to /a/ but still with almost identical ratings. In this study, AE /ɑ/ assimilates to PA long (/a:/ and /u:/) but with better assimilation to PA /a:/ than to /u:/, despite the low ratings for both of them (/ɑ/ is unclassified). I interpret these findings as evidence that AE /ɑ/ is phonetically perceived as a long (and tense) vowel by PA learners of AE.

3.4. Discussion

Through conducting a PAM test, the present study aimed to unravel how the 11 AE tense/lax monophthongs are assimilated to PA 6 long/short vowels and which AE vowels resist assimilation most. Since no clean one-to-one mapping is possible between AE vowels and PA vowels, it was expected that all AE monophthongs would be problematic for PA learners, especially in the mid-vowel section, due to lack of contrast with PA (see section 2.1.4.1 and Figure 2.3). The main findings with regard to the perceptual assimilation of the mid vowels show that PA participants have a good sense of the AE vowels, as their perceptual assimilation was only skeptical concerning spectrally adjacent vowels, and no confusion was observed in the front-back dimension. Since PA has no vowels in the mid-section, it was expected that learners would assimilate AE mid vowels to their nearest PA high vowels but that there would be no confusion in tenseness at all.

Within the PAM framework, the results are stratified according to their assimilation types. TC contrasts are the least demanding, followed by CG contrasts, and SC contrasts are the most demanding. The results revealed that, in general, the large majority of the possible AE vowel contrasts should not be problematic for PA learners. These were the contrasts categorized with a TC scenario (47 out of the 55 possible contrasts). Meanwhile, 7 possible contrasts were found to embody an SC scenario, and only one would be a CG contrast. Since TC categories impose no learning problem, these require no further discussion. Among the hypotheses on vowel assimilation (Figure 2.3), an exception was provisioned for a possible confusion between AE

/ʌ/ and /ɛ/ to assimilate to PA /a/, and if that would turn out to be the case, PA learners would be able to discriminate AE /ʌ/ from /ʊ/ by forming a TC contrast for them. This provision was not supported by the results; rather, the more general hypothesis was supported that AE /ʊ/ and /ʌ/ fall in an SC contrast. Eventually, PA /a/ was not an assimilation category for any AE vowel.

Although all of the eleven AE pure vowels were categorized by the PA participants, most of the difficulties in perceiving the AE vowel inventory were more likely in vowels that do not have a clear counterpart in PA than in vowels with (almost) direct equivalents. The vowel that was most difficult to assimilate to the PA vowel inventory was AE /ɑ/ with a 2.2 fit index as a token of PA /a:/ and 1.0 fit index as PA /u:/. Next comes AE /ɔ:/ with /3.1/ fit index with poor resemblance to PA /u:/.

3.4.1. PA long/short vowels compared to AE tense/lax vowels

The results demonstrate that AE monophthongs were mapped onto the PA vowel inventory in a systematic fashion in the sense that all AE lax vowels only assimilated to PA short vowels, while the AE tense vowels were assimilated to PA long vowels without any overlapping in regard to tense ~ lax features. More precisely, the participants showed degrees of success in perceiving AE vowels not only to be based on duration but also with a clear-cut categorization according to their qualitative features in such a way that is not vague or overlapped and supports the interpretation of the results. However, most of the time, the assimilation within each contrast set was not perfect, first, because of the mismatch between L1 and L2 and the interference of the L1 remnant in their L2 vowel inventory. Therefore, it is only natural for learners to confuse some adjacent vowels due to quality differences. Second, it seems that learners have not developed a perfect sense of AE vowel quality. Therefore, most of the AE monophthongs were categorized as SC scenarios and only one CG scenario without yielding any uncategorized vowel.

The present findings of this study, as summarized in Figure 3.2, can be compared to analyses and predictions of L1 interference as experienced by other Arabic learners of American English. Yavaş (2011) suggests that AE vowels assimilate to Arabic vowels in a different way, as shown in Figure 3.3 below.

Yavaş (2011) compared the vowel systems as shared between AE as L2 and many languages as L1, including Arabic. In his analogy, he presents the Arabic vowel system as a 3-vowel system that can be doubled only in duration, which is far smaller than that of AE, a notion that highlights the importance of the insufficient separation (i.e., under-differentiation) of L2 phonemic distinction by Arabic L1 learners. Accordingly, he depicted possible confusions that

Arabic L1 learners can fall victim to when acquiring AE. The frequently attested lack of contrast (i.e., homophony) between Arabic and AE, according to Yavaş (2011), is observed in the following groups of AE vowels: AE /i, ɪ/ (e.g., *heed-hid*) are expected to be a rendition of Arabic /i/, /ɪ, e, ɛ/ (e.g., *hid-hayed-head*) are expected to be a rendition of Arabic /i/ vowel, /ɛ, æ, ʌ, ɑ/ (as in *head-had-hud-hod*) are expected to be a rendition of Arabic /a/, and finally, high back vowels /u, ʊ/ (as in *who'd-hood*) are confused for Arabic /u/.[25] These target contrasts are claimed to be overlooked by Arabic L1 learners. Additionally, vowels such as /o/ and /ɔ/ are not classified at all, neither as not problematic at all nor outside the possible rendition of L1 Arabic learners of AE altogether. Moreover, there is an overlapping across boundaries in terms of tense/lax contrast for the vowels /ɪ/ and /ɛ/ to assimilate to more than one vowel in the Arabic vowel system.

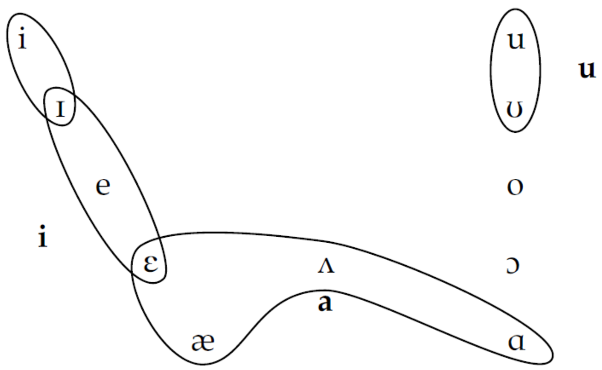

Figure 3.3. Arabic vowel assimilation over AE vowels. As in Yavaş (2011:197).

Figure 3.3 from Yavaş (2011) does not differentiate between long and short vowels. Yavaş suggests that AE /i:, ɪ/ (as in *heed-hid*) will be insufficiently distinguished in Arabic EFL, which would be incompatible with the perceptual assimilation results I reported above, where the AE pair was mapped on short PA /i/ and long /i:/, respectively. The same discrepancy can be observed in the case of the AE back vowels /u:, ʊ/, which were clearly separated in the assimilation results. Next, Yavaş suggests that Arabic EFL speakers do not properly separate the mid front vowels /ɪ, e:, ɛ/. The results of the present chapter are in alignment with Yavaş so far that indeed /ɪ/ and /ɛ/ map onto the same vowel in PA (i.e., short /i/) but /e:/ is not part of the confusion set because it differs in duration. Based on the assimilation results, it would be very unlikely that PA EFL learners would fail to make a distinction between /ɛ/ and /æ/ because these assimilate to different PA vowels, i.e., short /i/ and long /a:/, respectively. Yavaş does not suggest a problem with regard to the contrast between the lax vowels /ʌ, ʊ/, although both

[25] Arabic vowel symbols are as used in Yavaş (2011).

assimilate to PA short /u/ in the results. Finally, AE /o:/ and /ɔ:/ are not classified as problematic for Arabic learners of AE, although both vowels, together with AE /u:/ map onto the same vowel in PA, i.e., /u:/.

Briefly, my results for the PA learners in Figure 3.2 differ substantially from Yavaş (2011). I found one pervasive, clear-cut separation of the vowel categories: L2 lax vowels assimilated to short PA vowels, while AE tense vowels assimilated to long PA vowels. Further assimilation patterns were based on adjacency in vowel position in the IPA chart. The categorization that is suggested by the results of this study is more natural, with no obvious overlapping across categories. Vowel duration as part of the lax/tense contrast in AE will not impede PA learners in their acquisition of AE. In contrast, I predict that PA learners will additionally use quality differences in their identification of members of the AE tense/lax pairs.

These findings suggest that the PA vowel system is sensitive to vowel quality as well as to vowel duration, in line with Saadeh's (2011) results (in a production test), which showed that the PA vowel system centralizes the short vowels /i, u/ relative to their long counterparts /i:/ and /u:/ in a way that is similar to English [ɪ] and [ʊ] in AE. However, Saadeh's results cannot be considered conclusive, as her participants are heritage speakers of PA (i.e., second- and third-generation immigrants) and have not lived in Palestine or learned PA as L1. The results of perception studies of English vowels by listeners from different colloquial Arabic varieties also go in line with my results and show divergence from the classical Arabic vowel system, which is built on only three vowel features (height, constriction place, duration) toward additional features such as tense/lax or peripheral/non-peripheral (e.g., Fathi & Qassim, 2020 on Iraqi Arabic, Almbark, 2012 on Syrian Arabic, Nikolova-Simic, 2010 on Saudi Arabic, Watson, 2002 on Egyptian Arabic, Kalaldeh, 2018, on Jordanian Arabic, and Al-Mazrouei, Negm & Kulikov, in press, for Qatari Arabic).

3.4.2. Single-Category (SC) Contrast

Only 7 AE vowel pairs (out of 55) are in an SC contrast with one PA vowel. These contrasts occurred in adjacent vowels that only differ spectrally. In relation to PAM predictions, SC contrasts are predicted to be difficult to discriminate (Best, 1995; Tyler, 2019). In such contrasts, PA learners face strong demands toward conflating their perceptions for the confused members of each contrast. Therefore, their discrimination will be poor.

Referring back to the last questions of this study concerning the alignment of the results for PA and the results of studies on other L1 varieties, the way the PA participants map the large AE vowel inventory onto their smaller Arabic inventory is in line with results from previous studies

(Nikolova-Simic, 2010, Almbark, 2012; Faris et al., 2016; Evans & Alshangiti, 2018, Al-Abdely & Thai, 2016a, b). Hence, negative language transfer is expected, especially in the case of the single category assimilation pairs. However, my findings deviate from similar studies on Arabic L1 in regard to the question of which vowels are the most problematic for PA learners. Many of the previously mentioned studies (e.g., Nikolova-Simic, 2010; Al-Abdely & Thai, 2016a, 2016b; Evans & Alshangiti, 2018) have reported that /ɛ/ was (among) the most problematic vowels for Arabic learners of English within their L1 variety. In contradistinction to this, the PA results in the present study, however, show that AE /ɛ/ is slightly better assimilated to PA /i/ ('fair', with a 3.7 fit index) than AE /ɪ/ ('poor' with 3.5 fit index). This could be reason to place the /ɛ-ɪ/ pair in a CG contrast, which is less problematic than an SC scenario.

3.4.3. Category-Goodness (CG) Contrast

Concerning the question of which vowel(s) are predicted to be the most problematic to perceive for PA learners of EFL, only one CG contrast appeared in the PA results, i.e., for the /æ ~ ɑ/ contrast, as both assimilated to PA /a:/. Note also that AE /ɑ/ is a member of another SC contrast with /ɔ/. The fit index for AE /ɑ/ as PA /a:/ is 2.2, while its fit index as PA /u:/ is 1.0 in the contrast /æ ~ ɑ/ as they assimilated to PA in CG contrast. The results suggest that learners will soon be able to consider PA /a:/ a good exemplar and representative of AE /æ/ before a new vowel category will be developed for AE /ɑ/. The situation is almost parallel for the AE /ɔ ~ ɑ/ SC contrast. Even if they were within an SC contrast that presupposes that both vowels are at an equal distance from the PA /u:/ vowel, the fit index (3.1) for the AE /ɔ/ suggest that learners will consider it a better exemplar of PA /u:/ than of AE /ɑ/ with a fit index of 1.0.

3.5. Conclusion

The results show that PA learners confused all AE monophthongs in the sense that there is not a single case of a PA vowel that maps uniquely onto a single AE vowel. As shown in Table 3.2, every PA vowel is associated with at least two different AE vowels. The first question of this study asked how the eleven AE monophthongs assimilate to the six-member vowel inventory of Palestinian Arabic. The results show that there is a general confusion in mapping AE vowels due to the PA participants' unfamiliarity with the quality differences of the AE vowels, especially in the mid and back sections, which were the most confused. Additionally, there cannot be a one-to-one mapping between L1 and L2 because of the mismatch in the vowel inventories, which implies that learners cannot achieve better performance based only on the differences between the vowel inventories of the two languages. Therefore, without explicit teaching tasks and aids that focus on the problematic vowels to help reach a better mapping in the participants' L1, a better performance cannot be reached. The confusion of the PA vowels as assimilating category for AE monophthongs is reflected in PAM contrasts yielded from the results where two or more AE monophthongs assimilated to one of PA 6 vowels, except for the PA /a/ which never reached the threshold to be considered an assimilation category for any AE vowel. The most troublesome monophthong is expected to be the AE back low /ɑ/ vowel, as it received the lowest fit index among all AE vowels and fell in a CG scenario with other AE vowels: with AE /æ/ as an assimilation category for PA /u:/ in a CG contrast and in an SC contrast with AE /ɔ/ as they both assimilated to PA /u:/.

The second question was concerned with how the PA results align with studies of other Arabic varieties to discover whether there are shared L1 interference effects when perceiving English vowels. In addition, the third question asked how sensitive PA learners are toward AE vowels in terms of timbre (quality) and duration. The results revealed which vowels were the most difficult to assimilate by PA learners. The findings were in conformity with some L1 Arabic studies in regard to certain AE vowels, but they do not agree with other results. In conformity, my results show that PA participants' assimilation behavior reflects a sense of AE vowel length differences, as their perceptual mappings do not show any confusion in the duration feature (i.e., they assimilated AE lax vowels to PA short vowels and AE tense vowels to PA long vowels). My results also confirm the notion that the PA L1 inventory is not exclusively differentiated based on duration but also includes spectral differences between long and short counterparts, as found in Al-Ani (1970: 24), Watson (2002: 22) and many studies on L1 different Arabic varieties mentioned in section 3.7.1.

The PA learner's association with AE vowel qualities beyond duration was also demonstrated to some extent, especially in Figure 3.2. Moreover, my hypothesis that learners will be aware of the spectral differences between adjacent sounds but will rely primarily, in their assimilation judgments, on the durational cue to discriminate between AE vowels was confirmed. PA learners demonstrated sensitivity to other qualities of AE in addition to duration. The following chapters will provide data on PA discrimination and the production of AE vowels, which can be interpreted in light of the results of this study.

CHAPTER FOUR

PERCEPTUAL MAPPING OF THE AE VOWEL SPACE BY AMERICAN NATIVE LISTENERS AND BY PALESTINIAN-ARABIC LEARNERS OF ENGLISH AS A FOREIGN LANGUAGE

4.1. Introduction

4.1.1.Overview

In this chapter, I provide a comparative analysis of the perceptual representation of the vowel space of American English (AE) as conceived of by nonnative Palestinian Arabic (PA) learners of English as L2 (group A) and by native speakers of AE (group B) through an identification task with synthesized vowels. The study aims to uncover differences in the mental representation of the vowel space between native AE speakers and nonnative PA learners of AE. The main objective of this study is to determine whether the nonnative speakers perceive the AE vowel space the same way as the native AE speakers do and, if not, what their perceptual representation looks like (or sounds like) in terms of vowel color (vowel quality as determined by formant structure), duration (quantity) and the relative importance (trading relationship) between color and duration.

This chapter is organized as follows: Section 4.1 provides a general description and the background of the study to contextualize it in relation to the available literature. This section also presents the questions and hypotheses related to the experiment. Then, in Section 4.2, the employed methodology is described and detailed as implemented in the vowel identification task, highlighting the procedures and the actual steps of the test. Next, in Section 4.3, the results of the test are presented for each of the two groups. Section 4.4 contains the results of this experiment and compares the results of both groups of participants. Finally, Section 4.5 draws conclusions and discusses the findings and their implications.

4.1.2. Background of the study

In many crosslinguistic studies of vowel systems, the aim is generally to pinpoint the essence behind the organization of these systems in relation to the number of phonemes available and how they are located inside the articulatory, acoustic, or perceptual space (Hawks & Fourakis, 1995), especially for vowels that bear most of the accent/intelligibility of the words. Applying this process to the vowels of any language provides us with what is called the perceptual mapping of language-specific vowels in the universal human vowel space.

Basically, the idea behind the human vowel space is to locate the different vowel sounds in a multidimensional space defined by the first two vowel formants (resonances), e.g., plotting the vowels as points in the F1-by-F2 coordinate system. These coordinates are considered to be associated with specific postures of the articulatory organs in the speech tract, mainly the position of the tongue body in a plane defined by vowel height (i.e., openness) and location of the constriction point (i.e., the front-back dimension). Here, F1 corresponds to the degree of openness (jaw aperture): the more open the mouth (and the larger the distance from the body of the tongue to the palate), the shorter the effective length of the throat cavity, and the higher the frequency of the first formant (F1, or lowest resonance frequency in the vocal tract, e.g., Ladefoged & Disner, 2011). The second-lowest resonance frequency (F2 or second formant) is related to the length of the oral cavity, which is determined jointly by the location of the constriction point and by the protrusion of the lips.

The analytic problem becomes more complicated if the learner's native language has no length contrast in its vowel system, while the target language systematically employs duration to divide the vowels into two (or even three) categories, whether short vs. long or lax vs. tense. Especially in the case of a target language with a tense-lax opposition (where duration codetermines the category in interaction with spectral properties F1 and F2), trading relationships between vowel duration and quality will be found for native speakers that will be absent in the mental representation entertained by nonnative learners. Van Heuven et al. (2020) stressed the idea that in the context of teaching the phonetics of a foreign language, some parts of the vowel system organization may still be vague for learners, and more specifically, the perceptual representation of the vowel system in L2 is often incorrect. Van Heuven et al. (2020) also call for considering how the vowel systems are represented in the learner's mind (both native and nonnative) through a perceptual mapping technique. Employing such techniques can reveal trading relationships across the parameters that actualize each listener's vowel space.

4.1.3. Study questions

This chapter aims to compare how the native AE perception of vowels differs from the nonnative PA perception of the AE vowel space. In the previous chapter, I found several potentially problematic contrasts for the PA learners. More particularly, the results of the present chapter will be compared to the PAM task results from the previous chapter to either confirm or refute them. Additionally, by mapping out the perceptual representation of the nonnative vowel space, this chapter will help pinpoint which AE vowels are difficult to discriminate for PA learners by visualizing how AE vowels (or vowel coordinates) overlap one another in the mind of PA learners compared to the native AE results.

Concretely, the aim is to answer the following three questions:

Q1. How do Palestinian Arabic listeners conceive of the American English vowel space?

Q2. How do American native listeners conceive of their own vowel space?

Q3. How does the mental representation of vowel sounds differ between AE native listeners and PA learners in terms of vowel quality and duration, and how do the quality and the duration interact or trade?

4.1.4. Hypotheses

The main focus of this vowel identification task is to examine in what aspects the PA participants' perception of the AE vowel space differs from that of the natives. My hypotheses, in this case, are based on the results of the PAM task and the vowel contrasts that PA learners found problematic based on the assimilation patterns and whether they are the only problematic ones or new difficulties can be found in the identification task in the present chapter.

Additionally, even if English (in general) and Arabic have (near) equivalent vowels, I hypothesize that even for those *shared* vowels, their perception within the AE vowel system for the nonnative PA learners will differ from the native's because they differ in acoustics cross-linguistically. Boundaries between adjacent vowel categories will differ between PA and AE listeners. The vowels that are shared between PA and AE (short/long /i/ and /u/ will show up as expanded /i/ and /u/ areas in L2, while the target vowels in the mid-section of the vowel space will shrink, i.e., will take up less space than they do in native English. Additionally, the nonhigh vowels, such as in *bed*, *bad*, *bud*, *bawd* and *bod*, will not be good exemplars of any PA vowel, so they will be poorly defined in the learner's perceptual representation of the target AE vowels. Finally, there are the tense mid vowels as in *bade* and *bode*. In AE, these are not just long but also diphthongized to a certain extent. The long monophthong versions in the synthesized vowel tokens (see below) may not be very acceptable to native AE listeners, but quite possibly PA learners of English will accept these long monophthongs as adequate tokens of the target vowels.

Therefore, for the present study, I hypothesize the following:

H1. PA learners' perception (or subdivision) of the AE vowel space is generally incorrect, except for the AE high vowels.

H2. PA learners' perception of the AE vowel space will reflect the organization of the PA L1 vowel system.

H3. The trading between AE vowel parameters in the mind of PA learners will only reflect the features available in their L1, i.e., a rough tripartite division of the vowel color space PA vowels, subdivided further by differences in duration.

4.2. Method

4.2.1. Preliminaries

Among the different techniques available for creating/synthesizing a vowel space, a uniform systematic sampling of the complete vowel space can be obtained using an acoustic vowel triangle that is defined by F1 and F2 as independent parameters, while higher formants can be fixed or correlated with F2. Vowel duration can be manipulated as needed (if relevant in the target language). The first two formants are incorporated as patterns yielding a synthetic token. Learners (and native listeners of the target language) are asked to identify each token as a vowel of a specific language (whether L1 or L2). This token and the coordinates it represents are assigned the phonetic symbol for the chosen vowel. Then, tokens labeled with the same vowel (same vowel symbol in the chart) are grouped together to create a cloud of measurement points that constitutes a subspace within the vowel space. The clouds can be defined by their centroid (located at mean F1 and mean F2 of the vowel type) and some dispersion measure, typically a spreading ellipse drawn at ± 1 standard deviation (SD) away from the centroid along the first two principal components of the scatter cloud, thereby including the 46% most central tokens of the vowel type (e.g., Klein et al., 1970). Sometimes, especially for the nonnatives, the distributions (and spreading ellipses) of adjacent vowels can overlap or become entangled, reflecting a misperception of the vowel space in the TL. This technique is followed in this chapter following Van Heuven et al. (2020) but in my case for AE native listeners and PA nonnative learners of EFL (rather than L1 Persian learners of English).

In this experiment, the listeners identified each stimulus as one of the 11 AE pure vowels to estimate the perceptual tolerance (in terms of modal response) and distribution (defined by the centroid and the dispersion) of each AE vowel type. These distributions allow me to define boundaries in the three-dimensional vowel space (defined by F1, F2, and duration) that optimally describe how the listeners discriminate between two adjacent vowel phonemes.

4.2.2. Stimulus material

The stimuli were a universal reference set of synthesized vowel tokens in a C_C carrier that allowed researchers to map out the perceptual representation of the pure vowels (monophthongs) of any language (Van Heuven et al., 2020). The proposed vowel space is defined by three independent parameters, i.e., vowel height (F1), vowel backness/rounding (F2), and length (vowel duration). The synthesized reference set of tokens contains a labial onset and coda, /m/ and /f/, respectively, and a synthesized nucleus in between them, yielding a /m_f/ carrier. The labial consonants can be easily synthesized and have approximately the same articulation in any language.[26]

The sampling of the synthesized vowels to create a universal reference vowel set was done by making equal steps (of 1 Bark) in a grid of 7 steps along the F1 dimension and 9 steps along the F2 dimension yielding 63 combinations, 20 of which are impossible (inhuman) and therefore excluded.[27] Therefore, this yields 43 vowel qualities that can be manipulated in duration if needed for any language. I opted for two vowel durations, i.e., 200 ms (short) and 300 ms (long), for all synthesized vowel tokens; 86 stimuli in total.[28] The stimuli were synthesized using the programmable LPC formant synthesizer implemented in Praat (11,025-Hz sampling rate, 16-bit amplitude resolution) and saved in mono (one-channel) .wav file format. Figure 4.1 shows the organization of the vowel space that was used in the experiment.[29]

[26] Not all languages, however, have /m/ and /f/. Here the user may want to change the /m/ to a /b/, and/or the /f/ to a /p/. This is not needed for English and Arabic.

[27] Either the F1 value exceeds that of the F2, so that the same vowel quality is synthesized as with F1 and F2 reversed, or the F1 is combined with an F2 value that cannot be produced by the human vocal tract.

[28] The vowels were synthesized in an acoustic vowel triangle marked by the F1-F2 values of extreme and prolonged vowels pronounced as /mif/, /maf/ and /muf/, technically not cardinal vowels but contexted tokens which are a very close approximation to the cardinals (Van Heuven et al., 2020). The extreme tokens synthesized were more peripheral than the human originals.

[29] The synthesis of the stimuli was done by Linear Predictive Coding as implemented in the Praat speech processing software (Boersma & Weenink, 2019). The script needed to synthesize the materials in the various spectral and temporal conditions was written by Jos J. A. Pacilly, software engineer at the Centre for Linguistics at Leiden University, The Netherlands.

F2 (across) F1(down)			F2 (step, Bk, Hz)								
			1.	2.	3.	4.	5.	6.	7.	8.	9.
			14.0	13.0	12.0	11.0	10.0	9.0	8.0	7.0	6.0
Step	Bark	Hertz	2357	2031	1746	1497	1278	1086	915	764	628
1.	2.5	237									
2.	3.5	339									
3.	4.5	447									
4.	5.5	565									
5.	6.5	694									
6.	7.5	838									
7.	8.5	998									

Figure 4.1. Steady-state F1 and F2 values for reference vowels. F1 is varied in 7 steps of 1 Bark (with equivalent hertz values shown), while F2 is varied in 9 steps. Twenty impossible/inhuman F1-F2 combinations are excluded, leaving a vowel triangle of 43 perceptually equidistant points.

4.2.3. Participants

Typically, participants involved in identification experiments are either native speakers of the language or phonetically untrained nonnative speakers. In this experiment, I worked with both types of participants to (1) differentiate between the native and the nonnative conception of the AE vowels in the mind of the participants two groups and (2) map out and compare their perceptions of sounds, which can reveal the effects of L1 in the results of the nonnative AE participants when compared to native AE speakers. Moreover, the AE native results will provide an optimal mapping of what the AE pure vowels should be like, which may serve as a model for any nonnative learners of AE.

The nonnative PA group of participants was the same group detailed in section 3.2.3 in the previous chapter (Group A in this experiment). The AE native listeners were 20 native AE (13 male, 7 female). Regarding their age, there were 6 elderly and 14 young adult native speakers of American English, hailing from eleven different states in the USA (nine from California). The young adults were students at the University of Southern California (USC) in Los Angeles.[30] More demographic details about the AE native listeners are available in Appendix 4.1.

4.2.4. Procedure

The procedure of this experiment was the same as in the previous experiment unless stated otherwise. The 86 synthesized stimuli described in section 4.2.1 (43 short tokens and another 43 long tokens) were made audible only once and one at a time in a different random order for

[30] The data in Van Heuven et al. (2020) were based on the first 16 AE participants. The results reported in the present chapter were based on four more native listeners. The data were made available for use in this chapter by professors Vincent J. van Heuven (University of Pannonia, Veszprém, Hungary) and Sandra F. Disner (University of Southern California, Los Angeles, USA).

each listener. The response buttons contain well-known and meaningful monosyllabic (real) English words written in English orthography in a fixed /CVl/ carrier covering the 11 pure vowels of AE. In the carrier, C refers to a consonant phoneme in the syllable onset position, V refers to the synthesized vowel (long or short in duration), and a fixed coda /l/. Response alternatives were the following words: *feel* /i/, *fill* /ɪ/, *sale* /e/, *tell* /ɛ/, *shall* /æ/, *null* /ʌ/, *doll* /ɑ/, *call* /ɔ/, *whole* /o/, *full* /ʊ/, and *fool* /u/ (see Figure 4.2). The participants were asked to click on the response button with the word that contained the vowel they had just heard. Participants were instructed to always click a button when they heard a stimulus and to gamble if they could not choose between alternatives. This renders the task an eleven-alternative forced-choice (11-AFC) identification task.

Then, participants were required to provide a judgment on a 3-point Likert scale of how good an exemplar of the selected vowel they considered the vowel they had just heard was (typicality judgment), where 3 = Good. The process was repeated until all 86 tokens were covered.

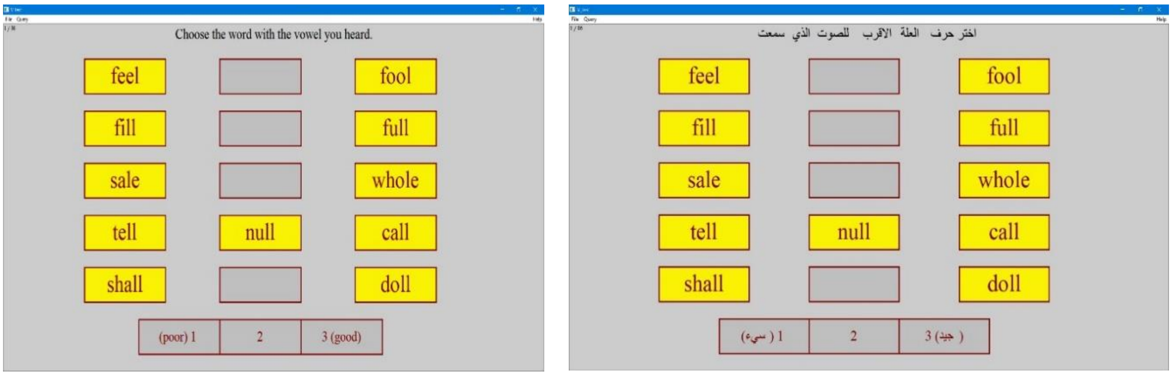

Figure 4.2. User interface for the vowel identification test for native AE speakers presented in English (left panel) and for L1 PA participants with instructions translated into Arabic (right panel).

Presentation of instructions and stimuli as well as the collection of the responses were performed by a Praat MFC script (see Appendix 4.2). The choices the participant made, the typicality judgments and the time it took the respondent to enter the typicality judgment (in milliseconds from the offset of the stimulus) were logged and stored for offline data analysis.

4.3. Results

4.3.1. Dividing up the vowel space

In this section, I will consider how the vowel space is divided up by the native and nonnative listeners. I will do this separately for the short vowel set and the long vowel set. Each set comprises 43 vowel points that only differ in their quality (vowel color). Figure 4.3 shows the category to which the stimulus was assigned for each sampling point in the vowel space. If 50% (or more, i.e., the majority) of the responses converged on a particular category, a large-print bolded phonetic symbol is entered in the figure. Agreement between 25% and 50% of the responses is indicated by a smaller print (unbolded). When less than 25% agreement was obtained for a particular vowel category, the cell was left blank (this rarely occurred). Therefore, the symbols in the cells of Figure 4.3 represent the modal response given to the synthesized vowel that is defined by the F1, F2, and duration coordinates defined by the cell position.[31] Figure 4.3A shows the mapping of the English vowel space as entertained by native American listeners responding to the short vowel stimuli, while Figure 4.3B shows the results of the long stimuli for the same group of participants.

A. Short stimulus vowels

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	**i**	**i**	u	**u**	**u**	**u**	**u**	**u**	**u**
2.	**i**	**i**	ɪ	ɪ	**u**	**u**	**u**	**u**	**u**
3.		ɪ	ɪ	ɪ	ʊ	**ʊ**	o	**o**	
4.		**ɛ**	**ɛ**	**ɛ**	ʌ	**ʌ**	ʌ	ɑ	
5.			**ɛ**	ɛ	ʌ	ʌ	ɔ		
6.				**æ**	**æ**	ɔ			
7.				**æ**	**æ**	ɔ			

B. Long stimulus vowels

1.	2.	3.	4.	5.	6.	7.	8.	9.
i	**i**	u	**u**	**u**	**u**	**u**	**u**	**u**
i	**i**	ɪ	ɪ	**u**	**u**	**u**	**u**	**u**
	ɛ	**ɛ**	ɪ	ʊ	ʊ	o	**o**	
	ɛ	**ɛ**	ɛ	ʌ	ʌ	ɑ	ɑ	
		æ	**æ**	æ	**ɔ**	**ɑ**		
			æ	**æ**	**æ**			
			æ	æ	æ			

X: >50% agreement; x: 25-50% agreement

Figure 4.3. Modal responses by 20 American native listeners for 43 vowel stimuli differing in F1 (vertically) and F2 (horizontally) center frequencies. Vowel duration is either 200 ms (panel A) or 300 ms (panel B). For specifications of F1 and F2 steps, see Figure 4.1. Large bolded symbols denote a majority decision with 50% or more agreement. Small symbols indicate a modal response with an agreement between 25 and 50%. Cells with a modal response < 25% agreement are left blank.

The native listeners have a rather straightforward division of their vowel space. The top left of the space is taken up by tense /i/, while the top center and top right areas are exclusively occupied by tense /u/. Then, going down along the front edge of the diagram, there is a well-demarcated area for /ɛ/ and a much smaller area for maximally open /æ/. The central portion of the vowel space is taken up by the other lax vowels /ɪ, ʊ, ʌ/. The open and half open back vowels /ɔ/ and /ɑ/ are not delineated from each other; in fact, open /ɑ/ is seen as an island surrounded by /ɔ/ responses. Note that all tense vowel categories – with the exception of the semi diphthong /e/ – are a preferred response for at least two sample points despite their short duration. This is an indication that duration is unlikely to be the primary cue in the tense-lax contrast in (American) English, in contradistinction to what was claimed by Wells (1982: 120).

[31] Occasionally, the distribution of responses to a particular vowel type was bimodal or even multimodal. This happened with 5 out of the 86 vowel stimuli for the native listeners, but never for the PA listeners. In such cases the modal response was assigned such that contiguous areas in Figure 4.3 were maximized. Appendix 4.3.A contains a complete breakdown of response category by stimulus type for the native participants.

Figure 4.3B shows the results in terms of the preferred responses of the American listeners for the 43 vowel types with long duration (300 ms). It would appear that category boundaries between pairs of vowels that differ in height are very regular and are almost exclusively based on F1 frequency. The tense vowel categories have expanded their size somewhat (from 26 to 31 sample points) since the vowel duration of 300 ms fits the specification of this type of category. By the same token, the area taken up by lax vowels has shrunk (from 17 to 12 sample points). The effect of longer duration is surprisingly large for /æ/, which vowel expands its area from 4 sample points to 9. This would be yet another indication that /æ/ in American English is phonetically tense and long. The semi-diphthong /e/ is never perceived with at least 25% agreement, not even when the duration matches its internal (long) specification. This would indicate that the closing gesture (diphthongal trajectory defined by gradual lowering of F1 frequency) is indispensable for /e/ but not necessarily so for /o/.

Now turning to the nonnative results, we may ask how PA learners, with their small L1 vowel inventory, map out the AE vowel space. The vowel labeling by the PA learners of English is shown in Figure 4.4A for short stimuli and 4.4B for long stimuli.

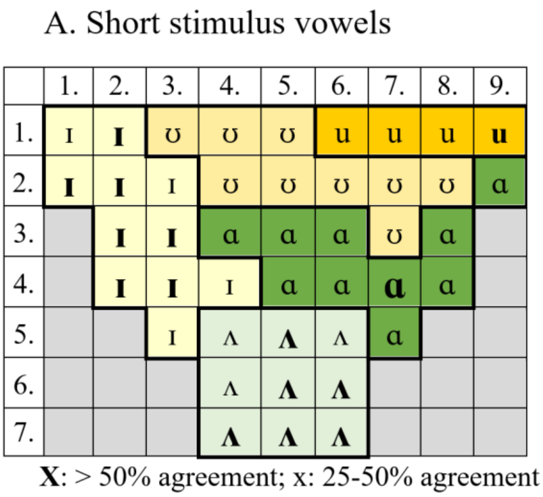

Figure 4.4. Modal responses by Palestinian listeners (20 boys, 20 girls) for 43 vowel stimuli differing in F1 (vertically) and F2 (horizontally) center frequencies. Vowel duration is either 200 ms (panel A) or 300 ms (panel B). For specifications of F1 and F2 steps, see Figure 4.1. Large bold symbols denote a majority decision with more than 50% agreement. Small symbols indicate a modal response with an agreement between 25 and 50%. Cells with a modal response < 25% agreement are left blank.

What is immediately noted for the nonnative PA listeners is that when the stimulus is short, the responses are mainly short/lax English vowels (see Figure 4.4A). Only 14 out of 43 stimuli were identified as long/tense AE vowels. Conversely, when the stimuli are long (300 ms), the responses are mainly tense vowels (see Figure 4.4B). Only 7 of the 43 stimuli were identified as short/lax vowels (exclusively /ʌ/).

The results for the short stimuli show mainly short/lax AE responses and therefore differ substantially from the native AE results on these stimuli (compare with Figure 4.3A). Lax /ɪ, ʊ, ʌ/ vowels obtain a majority vote only if the vowel duration is short (200 ms); they never reach a majority when the stimulus vowel is long (300 ms). Lax /ɛ/ never reaches 25% agreement, which would indicate that the mid vowels are indeed a special problem for PA learners of English. Duration, on the other hand, would seem to be a strong component of the (incorrect) perceptual representation of the AE lax-tense (short-long) contrast.

The top left area in Figure 4.4A is filled by lax /ɪ/, which overlaps and masks /ɛ/ downward, /i/ upward, and fills even part of the /æ/ area in the native AE listeners' results in Figure 4.3A. The top-right and top-central areas are fuzzily demarcated between a few /u/ tokens centered in the extreme upper right corner and an overgeneralized dispersion of /ʊ/ without any results >50% while occupying large chunks of what should be /u/ tokens. Similarly, /ɑ/ tokens spread horizontally across the triangle, which happens by shifting the tongue position from back to front but is still considered by the majority (25%-50%) of the participants as good representatives of AE /ɑ/; vowel fronting seems to be compatible with the perceptual representation of /ɑ/ on the part of the PA learners. The bottom part of the triangle is occupied solely by /ʌ/ tokens, while the other AE low vowels /æ, ɑ/ never reach the 25% agreement criterion. As shown in Figure 4.4A above, lax /ɛ/ is never a response category; it is always crowded out by /ɪ/, which means that PA learners of English do not have a proper separation of /ɪ, ɛ/ vowels. Because Arabic has no mid-vowel phonemes, AE short (lax) /ɛ/ is not conceived of as distinct from PA short /i/, which leads to the provisional conclusion that short front-mid vowels are considered allophonic tokens of the PA short high front vowel type.

It is possible to conclude from the results in Figure 4.4A that short stimuli mainly yielded lax responses, except for tense /u, ɑ/, which occupy undisputed areas along the back edge of the triangle. The rest of the tense vowels, especially /e, o/, never made it as a preferred ('modal') response category for the short stimuli, which indicates that duration is more likely to be the primary cue in the tense-lax contrast in (American) English in the perception of PA learners. The mid-central lax /ʌ/ is perceived too far low; low-mid lax /ɛ/ and the mid tense vowels never reach 25% agreement, which illustrates the problem with the mid-section vowels for L1 Arabic learners.

For the long stimuli, the results for the PA learners (Figure 4.4B) show that only /u/ and /i/ are well represented compared to other vowel tokens, but they are expanded substantially (relative to the native AE representation) when the stimulus duration is long (300 ms). All other long tokens are compromised when compared to the AE native results. The areas of /ʌ/ and /æ/

are about right – although /ʌ/ extends too far low and takes over sample points that should be /æ/, which in turn has shrunk to two sample points, and even these are too far back compared to the native norm.

It should be highlighted here that the longer stimuli yielded almost exclusively tense vowels as response categories. Accordingly, it is possible to say that duration is the most important cue for the PA EFL learners to differentiate between AE tense and lax high vowels. Duration, therefore, is given too much weight in the perceptual representation entertained by PA learners of American English. What confirms this claim in light of the AE tense/lax feature is that the presence of the AE lax counterpart ousted the tense counterparts and vice versa based on the stimulus duration, especially for the high vowels. Responses for the short stimuli filled 11 sample points with lax /ɪ/ in the high front area of the triangle, which covered the area that was supposed to be demarcated for tense /i/ and even /e/.

The opposite is true for the long stimuli where tense /i/ is fairly represented with the exact same 4 sample points as in the native results, and /e/ is never disputed with /i/. The same situation is applicable for AE /ʊ/ and /u/, but with better representation for the latter in the short stimuli than the former in the long stimuli. Additionally, /e/ and /o/ are represented as tense/long monophthongs, and diphthongization does not seem to be part of the perceptual representation of the tense English mid vowels in the mind of the Arab learners. Categories /ɑ/ and /ɔ/ are conflated, as they are in the vowel system of the Californian (and many other) native listeners. The difference from the native AE representation is that the AE vowels are long and tense; they never extend into the central area of the vowel space. In the perceptual representation of the PA listeners, the undifferentiated /ɑ, ɔ/ vowel pair is the short counterpart of tense/long /o/; both mid-central and mid-back vowel qualities are deemed adequate for /ɑ/ and /o/. All open vowels are perceived as /ʌ/, unless the vowel is long and very front, in which case /æ/ is perceived. Tense mid vowels /e/ and /o/ take up large areas in the vowel space when the stimulus duration is long. Apparently, the EFL learners accept these vowels on the basis of length and do not mind the absence of diphthongization. The AE vowel /æ/, which is long and generally considered phonetically tense in AE (Celce-Murcia et al. 2010; Strange et al., 2004; Wang & Van Heuven, 2006), appeared as the response for long stimuli because it resembles (near-equivalent) low front long /a:/ in PA.

To conclude this section, it seems as if PA learners of EFL have some notion of mid vowels but only when they are long. The short mid vowels are unified with their short high counterparts. The long mid vowels /e, o/ are entertained as viable response alternatives, even if they are unacceptable to native AE listeners because the PA learners do not know that the AE tense mid vowels should diphthongize.

4.3.2. Perceptual representation: Centroids and dispersion ellipses

Comparing the native and nonnative vowel identification results as shown above provides a systematic comparison for each parameter and component of the stimuli and how they are overlaid onto the vowel triangle. However, providing a side-by-side comparison using a pictorial display can show more about how exactly vowels/stimuli overlap each other and how their dispersion is superimposed. Figure 4.5 shows the native AE results.

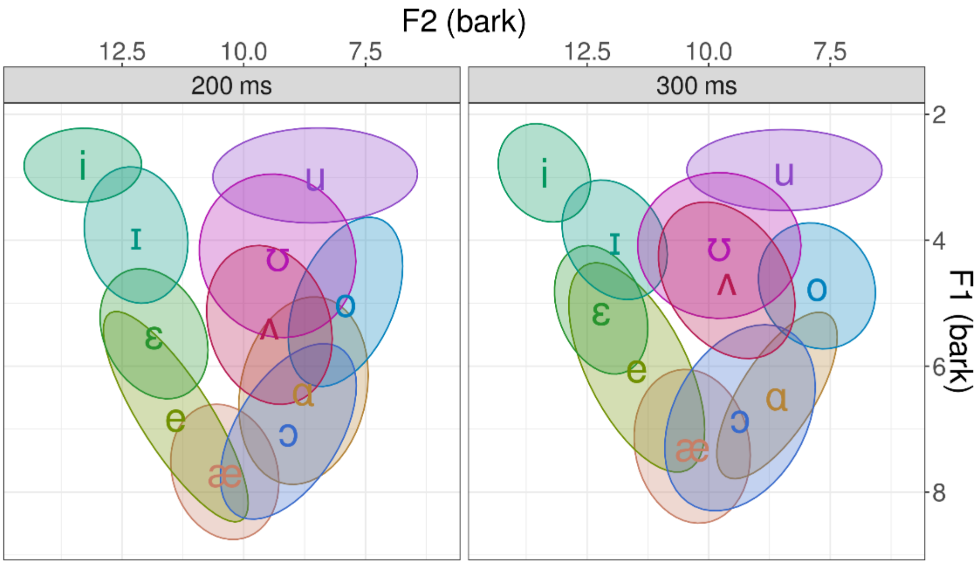

Figure 4.5. Centroids and dispersion ellipses (± 1 SD) in an F1-by-F2 plane (axes in Bark) for short (panel A) and long (panel B) stimulus vowels, as perceptually labeled by American native listeners. Phonetic symbols are placed at the category centroid. Spreading ellipses are drawn at ± 1 SD along the first two principal components of the scatter clouds (and theoretically include the central-most 46% of the data points per category).

On the other hand, Figure 4.6 shows the centroids and dispersion ellipses of the data in the F1-by-F2 plot based on the 20 male and 20 female PA participants. Data have not been weighed by typicality (i.e., all responses count equally). This technique provides very detailed insights that ease the interpretation of the results. Just with a cursory look, one can see how the centroids and dispersion ellipses for the PA participants in Figure 4.6 diverge from the native AE results in Figure 4.5.[32]

[32] The ellipse plots were produced with the Visible Vowels facility (Heeringa & Van de Velde, 2018).

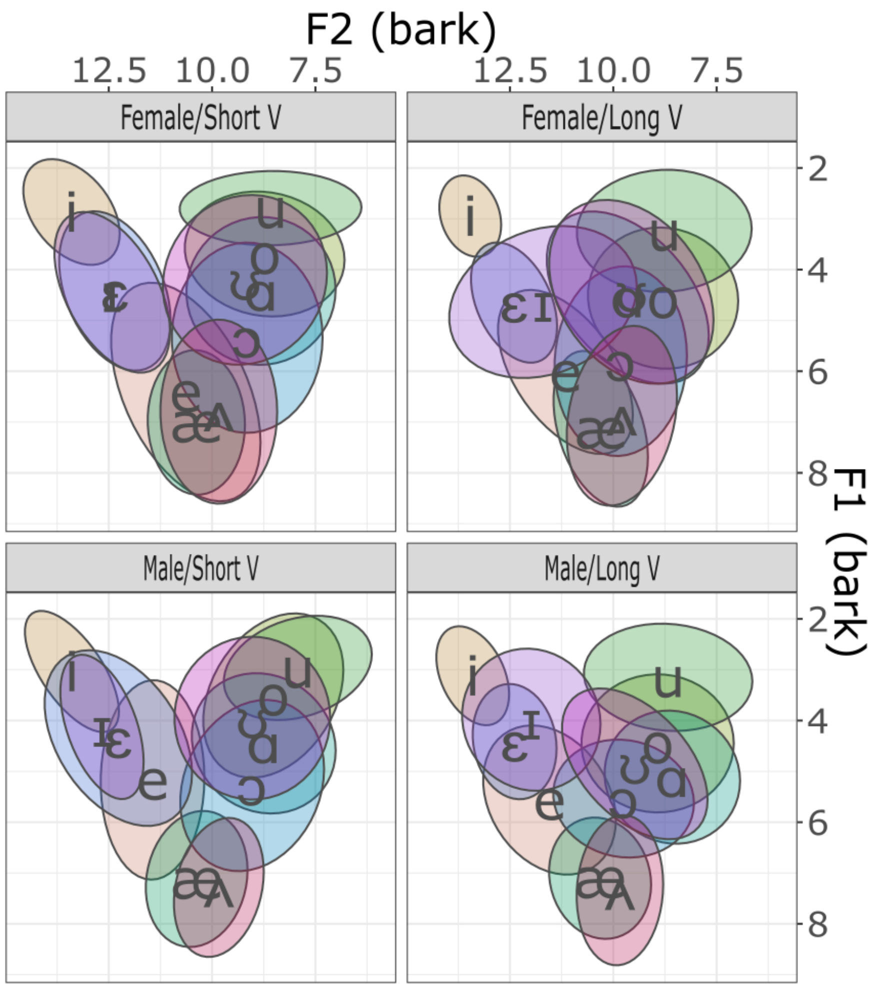

Figure 4.6. Perceptual labeling of synthesized short and long vowel sounds in /m_f/ context in terms of American English monophthong categories. Listeners were 20 male and 20 female adolescent speakers of Palestinian Arabic. For more explanation, see Figure 4.5.

4.3.3. Comparing the AE and PA results

To determine whether gender has any effect on the nonnative participants' accuracy in perceiving the AE vowel space in the synthesized set. My rationale for separating the results by gender at this point is motivated by the available research on how males and females differ with regard to their physical auditory systems and mental phonological processing (Krizman et al., 2012; McFadden, 1998) as well as by the finding of the LBQ questionnaire since the female EFL learners claim to have much better pronunciation skills in English than the men, and are more strongly motivated to pronounce English correctly and consider themselves better able than the men to perceive pronunciation details in a foreign language (See Table 3.1). One can argue that these differences are more salient in speech production, as males and females have different formant values caused by physical differences, but not in speech perception, where listeners do not impose their own voice on the test sounds as a frame of reference but rather use the context of the stimuli as the reference frame. However, previous studies have approached this attribute and stressed the importance of comparing male and female auditory perception in their natural settings (Eliot, 2013) and how crucial the differences are in perception based on gender differences in rendering adequate production, as proven by different studies (Almbarak, 2014; Flege 1991, 1995). The added value here is that I am inspecting the same factors but for perceiving synthesized sounds. It can also be connected to the participant performance and adequate production of speech sounds, as in the vowel production test in Chapter 5 of this dissertation. Accordingly, Figure 4.6 shows the parameters of the synthesized stimuli as they were interpreted by the PA learners of English broken down by gender.

In Figure 4.6, PA participants correctly placed /i/-, /u/- and /æ/-like stimuli at the left-top, right-top and mid-bottom parts of the F1-by-F2-plane, respectively. These vowel qualities were correctly interpreted without any problem given that they have near-equivalents in the Arabic vowel system. Other vowels, however, were scattered diffusely over the vowel space. Some differences in results based on gender of the listener can also be noted but without striking disparity, except maybe for /ʊ, ɑ/ for the long stimuli, and /ɪ, ɛ/ for the short stimuli in the female results (top panels in Figure 4.4), where the members of each pair of stimuli almost perfectly overlapped each other in centroids and dispersion, same as for the male results in the /æ-ʌ/ pair regardless of duration (bottom panels in Figure 4.4). Also, /ɪ/ and /ɛ/ are virtually indistinguishable in the male short stimuli. The /ɪ/ dispersion is engulfed by the /ɛ/ dispersion, and the centroids hardly differ.

What is noted for the PA learners (both males and females) is that they were able to allocate different profiles to vowels such as /ɪ~i/ and /ʊ~u/. Other vowel stimuli are well demarcated with appreciable spectral distances between centroids but partial overlap between their dispersions. Most likely, learners are in the developing phase, where they are trying to dissect and reorganize their perceptions to fit the native AE vowel space. Such claims can be confirmed via the production tasks in the following chapters.

Across panels, there are no striking differences in the positions of the vowel centroids except perhaps for the back vowels for females and low vowels for males. However, there is some shifting in the stimulus dispersions across panels, which can be connected to the differences in the duration of the stimuli.

The results of the vowel identification study can also be analyzed with respect to the importance of vowel duration in the mental representation of the AE vowels entertained by native AE and nonnative PA participants. One way to do this is to compute the mean duration for each of the eleven response vowels separately for the native and nonnative listeners. If it is true that the PA EFL students use the duration cue more than the native AE listeners do, one would expect to find a larger effect of response vowel on the mean vowel duration for the nonnatives. Figure 4.7 plots the mean vowel duration of the synthesized vowel tokens that were identified as instances of each of 11 response categories separately for the American native listeners and the PA EFL learners.

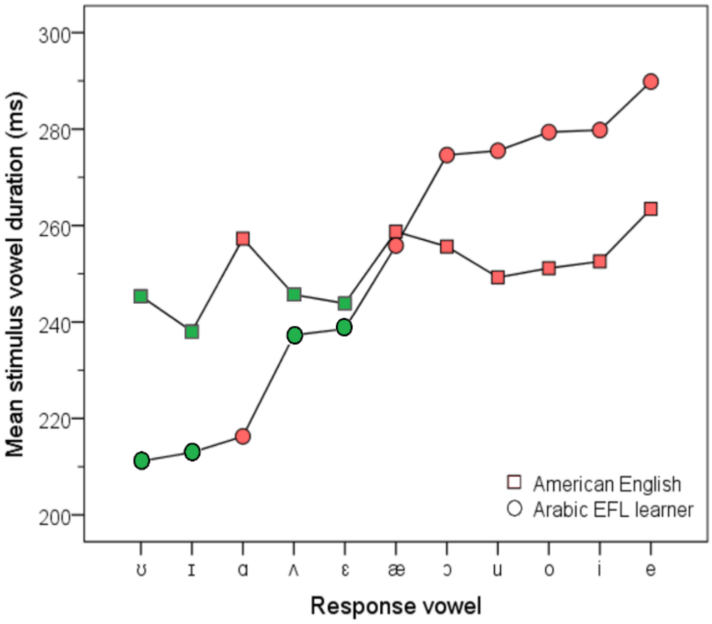

Figure 4.7. Mean stimulus vowel duration (ms) for each of 11 American English response vowels in the vowel identification task plotted separately for 20 American native listeners (square markers) and 41 Palestinian Arabic EFL learners (circles). Vowels are ordered horizontally in ascending duration chosen by EFL learners. Red markers represent phonetically tense AE vowels, and green markers represent phonetically lax vowels.

The effect of the response vowel in Figure 4.7 is much larger for the PA EFL learners than for the AE native listeners. The mean vowel durations run between 211 and 287 ms for the EFL learners, while they spread between 234 and 264 ms for the native listeners. A Repeated Measures Analysis of Variance (RM-ANOVA) on the mean durations per response vowel category per listener, with Native language (PA, AE) as a between-subjects factor and Response vowel as a within-subjects factor, shows that there is no main effect of Native language (251 ms for AE, 252 ms for PA, $F(1, 58) = .3$ ($p = .600$, p$\eta^2 = .005$). The main effect of Response vowel is highly significant, $F(10, 590) = 28.9$ ($\varepsilon = .649$, $p < .001$, p$\eta^2 = .333$), and so is the interaction between Native language and Response vowel, $F(10, 590) = 12.2$ ($p < .001$, p$\eta^2 = .173$).[33] Follow-up one-way RM-ANOVAs show that the effect of response vowel is significant for both native listeners, $F(10, 190) = 3.7$ ($\varepsilon = .549$, $p = .003$, p$\eta^2 = .163$), and for the PA EFL

[33] In these and all following RM-ANOVAs the assumption of sphericity was generally violated. Therefore, degrees of freedom were adjusted by the (GG) correction. The nominal degrees of freedom are multiplied by the epsilon coefficient ε (derived from Mauchly's *W*) mentioned in the text. I will report the nominal degrees of freedom but *p* values were determined after GG correction.

learners, $F(10, 390) = 71.7$ ($\varepsilon = .456$, $p < .001$, p$\eta^2 = .648$). The partial eta square values show that the effect size of the response vowel on duration is approximately four times larger for the nonnatives than for the native AE listeners. It can also be observed that the four short/lax vowels (green squares) have the shortest durations in the L1 data (< 250 ms), while the seven long/tense vowels (red squares) are all > 250 ms. The same criterion works for the PA vowel durations, with the exception of /ɔ/, which does not differ from the shortest lax vowels. The mean vowel durations per response category as selected by the natives and nonnatives are correlated, r = .585 (N = 11, p = .029, one-tailed), which indicates that the nonnatives have learned that vowel durations in English differ. However, it would appear that the English vowels /ɪ, ʊ, ɑ/ are conceived of as the short point vowels in their native language and that the vowels /ɔ, o, u, i, e/ are given the duration of long vowels in PA. The AE vowels /ʌ, ɛ, æ/, which do not occur in the learners' native language, are ambivalent in terms of the perceptual representation of the EFL learners: their mean durations cluster in the middle of the range between the short and long vowel categories.

4.3.4. Native and nonnative vowel identification compared in detail

Figure 4.4A-B shows the majority ('modal') decision for each of the 86 synthetic vowels in the stimulus material, as made by the PA EFL learners, while Figure 4.3A-B shows the native AE results for the same test. Quantifying the differences between native AE listeners and the nonnative PA students can be done by using the majority decision of the native AE listeners as the absolute criterion for the correct vowel identification. Using this majority criterion, many decision responses given by AE listeners will deviate from the majority decision and can be considered "wrong" or "suboptimal" responses or "confusions". The same criterion can be used to analyze the nonnative responses given to the same stimuli. In that case, the number of confusions will be considerably larger, and the discrepancies between the error responses for the two groups of listeners can be quantified by subtraction.

The first step in the analysis was to count the number of responses given to each of the eleven response categories (in the columns) for each of the 86 stimulus types (in the rows). Appendices 4.3.A-B contain the results of the counting for 20 American native listeners and 41 PA EFL students, respectively. On the left side of the table in Appendices 4.3.A-B, the responses are given for short vowel stimuli (vowel duration = 200 ms), and on the right-hand side shows the responses to the long vowel stimulus (300 ms). The rows are ordered by ascending F1 frequency and then by descending F2 frequency. The modal response for each stimulus type was then determined and is indicated in the table in boldface in green-shaded cells. In the native AE results, five stimulus vowels yielded a multimodal distribution of responses. In those cases, one of the multiple modes was given preference such that contiguity in the dispersion areas in Figures 4.2-3 was maximized.

If I take the modal response category of the group of native listeners as the correct or preferred response for a stimulus vowel, a percentage of correct scores can be computed for the American native speakers. Table 4.1 is a confusion matrix with the modal (preferred) category in the rows and the observed responses in the columns. The cells along the main diagonal contain the 'correct' responses for the native listeners. The off-diagonal cells contain alternative non-modal choices made, which can be considered 'wrong' or atypical.

Table 4.1. Confusion matrix of all observed responses against modal ('correct') response category for 20 American native listeners. Cells contain row percentages. Correct responses (agreeing with the modal response) are in boldface in green-shaded cells. Confusions ≥ 10% are in red-shaded cells. Marginals specify the number of observations in a row or column.

	All observed responses											Total
Modal response category	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	
i	**83.1**	11.3	1.9	2.5						.6	.6	160
ɪ	6.7	**42.5**		14.2	3.3		4.2		.8	11.7	16.7	120
e												0
ɛ	1.3	11.3	10.0	**67.5**	1.3	0.6	2.5	1.3		4.4		160
æ	.3	.7	9.0	12.7	**48.3**	8.7	2.0	18.0	.3			300
ɑ					5.0	**41.3**	6.3	20.0	25.0	1.3	1.3	80
ʌ		.7		5.0	3.6	12.9	**39.3**	5.0	7.9	25.0	.7	140
ɔ			2.5		20.0	23.8	2.5	**46.3**	3.8	1.3		80
o						3.8	13.8	7.5	**52.5**	20.0	2.5	80
ʊ		10.8	3.3	7.5	.8	.8	21.7	.8	.8	**41.7**	11.7	120
u	2.5	2.1		.8		.4	2.7	.2	1.5	12.1	**77.7**	480
Total	156	113	52	187	177	103	127	124	86	183	412	1720

The proportion of 'correct' responses (in cells along the main diagonal) is 1027/1720 = 59.7%. The vowels /e/ and /o/ are used quite infrequently as response categories. This will be due to the circumstance that the mid-high tense vowels in American English should be diphthongized. Moreover, /e/ never ends up as a modal response for any of the 86 synthesized vowels, so that the row for /e/ remains empty.

For the nonnative results, Table 4.2 (next page) repeats the same procedure for the PA students' responses to the same stimuli.

In the PA responses, the vowels /ɛ/ and /ɔ/ are used less often than the other nine categories. In fact, they are so infrequent as a response that they never end up as a modal response, so that two rows in the matrix remain empty. The proportion of modal responses in the PA relative to the total number of responses is 1828/3526 = 51.8%, so the conclusion follows that the PA listeners as a group are internally more divided in their responses than the native listeners.

Table 4.2. Confusion matrix of all observed responses against modal response category for 41 Palestinian Arabic learners of English listeners as a foreign language. For more information, see Table 4.1.

		All observed responses											Total
		i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	
Modal response category	i	**83.5**	3.0	4.3	5.5	.6	1.2	.6		.6	.6		164
	ɪ	7.8	**58.5**	2.7	16.4	2.7	2.9	2.9	.2	.4	4.9	.7	451
	e	4.9	9.8	**50.8**	17.1	4.5	.4	1.2	5.7	3.7	1.2	.8	246
	ɛ												0
	æ		1.2	24.4		**42.7**		18.3	8.5	3.7		1.2	82
	ɑ		1.5	1.0	2.0	3.9	**40.5**	13.2	6.3	2.4	27.8	1.5	205
	ʌ	.2	1.8	9.1	1.1	23.2	3.0	**48.8**	5.2	2.1	4.9	.6	656
	ɔ												0
	o	.5	.7	6.3	1.2	2.2	9.0	3.2	11.5	**43.4**	5.4	16.6	410
	ʊ	.9	4.2	1.0	3.0	.5	26.8	2.1	2.4	8.4	**36.8**	13.9	574
	u	1.6	1.4	1.1	1.1	.4	5.3	1.8	1.4	13.1	8.5	**64.4**	738
Total		204	346	266	166	234	349	417	140	357	411	636	3526

Table 4.3 presents the confusion matrix that is obtained when I compute the deviations in the PA responses from the norm that I derived from the American native data. These confusions would indicate where the perceptual representation of the American English vowels entertained by the PA EFL learners deviates from the native norm.

Table 4.3. Confusion matrix of all observed responses by 41 Palestinian Arabic learners of English as a foreign language against the modal ('correct') response category of 20 American native listeners. For more information, see Table 4.1.

		All observed responses											Total
		i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	
Modal response category	i	**51.4**	31.2	2.1	11.3	.3	.6	.9		.9	.9	.3	327
	ɪ	3.3	**34.1**	2.4	9.8	1.6	7.3	2.0	1.2	7.7	13.8	16.7	246
	e												
	ɛ	4.0	27.4	32.6	**17.1**	2.7	2.1	2.4	4.3	2.7	3.7	.9	328
	æ	.2	5.2	15.8	3.4	**27.3**	1.5	36.3	5.0	1.8	2.9	.7	615
	ɑ		.6	4.3		2.4	**19.5**	5.5	9.8	33.5	14.0	10.4	164
	ʌ	.3	1.4	2.4	1.4	8.4	21.6	**22.3**	8.7	12.2	16.4	4.9	287
	ɔ			6.7	.6	11.6	12.8	49.4	**5.5**	3.7	9.8		164
	o		.6		1.2	1.2	22.0	.6	7.9	**33.5**	19.5	13.4	164
	ʊ	.8	4.1	6.1	4.1		24.4	2.0	6.9	21.5	**19.5**	10.6	246
	u	1.1	2.2	.9	1.1	.3	10.4	1.7	1.2	11.3	18.1	**51.6**	984
Total		204	346	266	166	234	349	417	140	357	411	636	3526

Agreement between the PA decisions and the modal AE decisions is found in 1092/3526 = 33.8% of the cases. Clearly, then, the EFL learner's perceptual representation of the American vowel systems departs strongly from the native norm. Only for /i/ and /u/ do the learners concur with the native listener norm in more than 50%.

To illustrate how the PA conception of the AE vowel space differs from that of the native AE listeners, the deviations from the AE norm in the PA matrix are captured in Figure 4.8B. These deviations can be compared with the deviations from the AE norm (defined by the modal response by native listeners) in the native AE matrix, which are illustrated in Figure 4.8A.

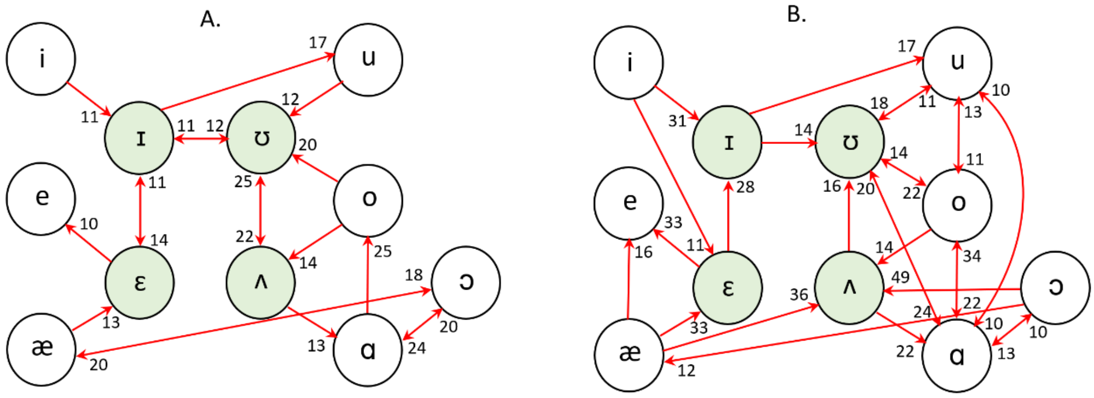

Figure 4.8. Vowel confusion structure of eleven American English monophthongs as identified for 86 synthesized vowel sounds by 20 American native listeners (panel A) and by 41 Palestinian Arabic EFL learners (panel B). Confusions < 10% have been omitted. Lax/short vowels in shaded circles. Arrows point away from the 'correct' modal vowel (according to the AE norm) to the incorrectly identified vowel. The percentage of confusion is indicated at the arrowheads.

The network in Figure 4.8 contains eleven nodes representing the monophthongs of American English. They are arranged in stylized fashion according to their position in a two-dimensional vowel space with vowel height (or F1) vertically and constriction place (backness, F2) horizontally. The seven phonetically tense vowels are on the outer perimeter, while the four lax vowels form an inner tetragon. The vowel /ɔ/ is positioned somewhat higher and to the right of low-back /ɑ/. This vowel seems to upset the symmetry in the AE vowel system, which may be one reason why /ɔ/ and /ɑ/ often merge into a single low back /ɑ/ vowel in American English. Single-headed arrows point away from the 'correct' (modal) vowel response to a nonmodal ('incorrect') response category. Only confusions larger than or equal to 10% are indicated. The percentage of confusion ('error percentage') is indicated at the arrowhead. Double-headed arrows indicate two-way confusion between two nodes.

Since the native listeners deviate from the modal response per vowel category in 40% of the cases, there is substantial confusion (or disagreement) even for the native listeners in Figure 4.8A. However, the disagreement is generally small and never in excess of 25%. Disagreement is largest for the vowel pairs /ɔ, ɑ/ and /ʌ, ʊ/, both of which are often implicated in current vowel mergers in American English.

In the PA results (Figure 4.8B), there is much more disagreement among the listeners, which indicates uncertainty on the part of the EFL learners regarding how to identify the vowels of

AE. In terms of the confusion network, there are many more arrows: 19 in Figure 4.8A (counting double-headed arrows twice) against 27 in Figure 4.8B, especially in the back vowel section. Some of the confusion is the same between the native and foreign listeners, such as the 17% confusion of /ɪ/ and /u/. However, the confusion and deviation from the AE modal response are substantially larger for the nonnative listeners in general, with the identification of /ɔ/ as /ʌ/ in 49% of the cases as the most notable one. Five of the confusion types occur in more than 25% of the cases. Interestingly, there are two pairs of vowels for which the nonnatives show less confusion than the natives. These are the low-back vowels /ɔ, ɑ/ and the lax vowels /ʌ, ʊ/. As mentioned above, however, the reason why the native listeners show substantial confusion for these pairs is that they are no longer distinguished by many American (especially Californian) English speakers. Figure 4.9 summarizes the differences between the confusion structure in the PA data and the native results. This figure shows the differences in the confusion found by subtracting the percentages in Table 4.1 from the corresponding cells in Table 4.2. Figure 4.9 thus shows the most important deviations in the mental conception of the AE vowels entertained by the PA EFL learners from the native norm, after correcting for uncertainty and deviations within the group of native listeners. What remains in Figure 4.9 is a relatively small number of confusions that deserve pedagogical attention because these errors in the perceptual representation are most likely to cause pronunciation errors. In the figure, deviations from the native norm (i.e., modal response) that occurred conspicuously more often with the PA learners than with the native listeners (a difference of 20 percentage points or more) are highlighted by thick red arrows. Negative numbers (and green arrows) identify discrepancies from the native norm, which should not negatively affect the learner's perceptual representation and pronunciation of the AE vowels.

Seven deviations (or errors in the perceptual representation) stand out and should require pedagogical attention. These are (i) tense /i/ is identified as lax /ɪ/, (ii) lax /ɛ/ is identified as tense /e/, (iii) tense /æ/ is identified as /ɛ/ (20%) or (iv) even as /ʌ/, (v) lax /ʊ/ is identified as tense /o/, (vi) tense /ɔ/ is identified as lax /ʌ/, and (vii) lax /ʊ/ and tense /ɑ/ are mutually confused. Without a single exception, these serious misrepresentations in the AE vowel system entertained by PA learners of EFL concern contrasts between AE vowels that involve a difference in the tense ~ lax feature. Typically, the EFL learners base their conception of the difference between the members of the compromised vowel pairs on duration, while they should attach (much) greater weight to the difference in vowel quality (including the moderate diphthongization that characterizes the tense mid vowels /e, o/ in American English).

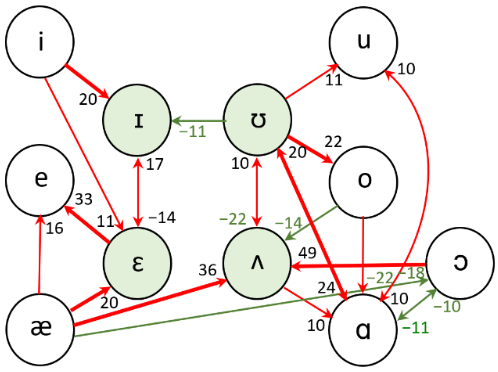

Figure 4.9. Difference in deviation (percentage points) from native norm (modal response category) in the perceptual representation of AE monophthongs entertained by Palestinian Arabic learners of English as a foreign language and by American native listeners. Thick arrows highlight the deviations from the native norm that deserve priority in EFL teaching. For more information, see Figure 4.8.

As a last illustration of the difference in weight attached to vowel duration in the mental representation of native listeners and PA EFL learners, Table 4.4 specifies how often each of the 11 vowels was identified by native and nonnative listeners across all 43 vowel qualities generated in the stimulus material, separately for short vowels and long vowels (for the complete breakdown of the responses by vowel quality and duration see Appendix A4.3.A-B for the native and nonnative listeners, respectively).

Table 4.4. Number of responses in each of 11 vowel categories to short vs long vowel duration in synthesized stimuli accumulated across all 43 vowel quality differences, broken down by language background of the listener (L1: American English native listener; L2: Palestinian Arabic learner of English as a foreign language. The absolute and relative differences in the number of responses are listed in the columns under Δ and %, respectively. Summary statistics are Chi-square and Phi. For more information, see the text.

		L1				L2			
		200 ms	300 ms	Δ	%	200 ms	300 ms	Δ	%
Response vowel	i	74	82	−8	−.5	41	163	−122	−3.5
	ɪ	70	43	27	1.6	302	44	258	7.3
	e	19	33	−14	−.8	28	238	−210	−6.0
	ɛ	105	82	23	1.3	101	65	36	1.0
	æ	73	104	−31	−1.8	104	130	−26	−.7
	ɑ	44	59	−15	−.9	292	57	235	6.7
	ʌ	69	58	11	.6	262	155	107	3.0
	ɔ	55	69	−14	−.8	39	101	−62	−1.8
	o	42	44	−2	−.1	74	283	−209	−5.9
	ʊ	100	83	17	1.0	363	48	315	8.9
	u	209	203	6	.3	157	479	−322	−9.1
	Total	860	860			1763	1763		
	All lax	344	266			1028	312		
	All tense	516	594			735	1451		
	$\chi^2(1)$	15.5 ($p < .001$)				617.1 ($p < .001$)			
	φ	.095				.418			

The Δ columns specify the difference in identification of the response vowel between the short and the long vowel durations. The %-columns express this difference as a percentage relative to the total number of responses given by the group of listeners, i.e., 1720 by the native listeners and 3526 by the PA EFL learners. It is then easily seen that the effect of vowel duration is much larger, also relatively, for the nonnatives than for the native listeners. In the table, the rows containing lax vowels are marked in yellow. The ratio of all lax vs. all tense vowel responses by native listeners is 40:60 for short vowels, against a ratio of 31:69 for long vowels. For the EFL learners, these ratios are 58:42 and 18:82, respectively. For both language groups, there is a significant association between tenseness and length of the vowels (long = tense, short = lax), but the association (expressed by the Phi coefficient) is much stronger for the EFL group (φ = .418) than for the native group (φ = .095). The difference between the two strengths of association is significant, $\chi^2(1) = 183.3$ ($p < .001$). This shows, once more, that duration is a (much) stronger cue in the tense-lax contrast in English for EFL learners than for native listeners.

4.4. Discussion

4.4.1. Overall accuracy of perceiving vowel quality

In the present study, I aimed to unravel the perceptual mapping of the AE vowel space for both nonnative PA EFL learners and native AE speakers and to compare their results in an identification task of synthesized long and short AE vowels. Regarding the first question, which was related to the overall accuracy of the nonnative PA learners in conceiving the AE vowel space, the results in Table 4.2 suggest that the PA learners struggle with their mental conception of the AE vowel space in general. The proportion of modal responses in the total number of possible responses was as small as 51.8%, while two AE vowels (/ɛ/ and /ɔ/) were never a modal response category for any stimulus vowel. For the second question of the present study, on the other hand, native speakers showed a better and more clearly defined perception of the set of stimuli, as their proportion of modal responses averaged at 59.7%, while native speakers also refused to perceive two specific vowels, viz. /e/ and /o/, which can be attributed to the lack of diphthongization of the synthesized stimuli set. This shows that diphthongization of /e/ and /o/, as in native AE human speech for the mid-high vowels, is a necessary prerequisite for their correct identification. Finally, the differences between the nonnative PA results and the native AE results were quantified to show how the PA perception of AE vowels departs from the native AE norm. This final step showed that the agreement of the PA decisions with the modal AE decisions is only 33.8%. At the level of the single response category, agreement with the

AE norm exceeded 50% only for two response categories, namely, /i/ (51.4%) and /u/ (51.6%). Agreement with the native modal response was much lower for the other nine vowel types and never reached 35%. This is strong evidence that the PA EFL learner's perceptual mapping of the AE vowel space is significantly different from that of the AE natives.

4.4.2. Overall accuracy of perceiving vowel duration

The results of the vowel identification experiment reveal that PA EFL learners have a fairly good perceptual representation of the duration of the vowels of American English. Figure 4.7 shows that the EFL learners assume short durations for the four lax vowels of AE and choose longer durations for the seven tense vowels. The correlation of the EFL and native durations preferred for each of the eleven vowel types is high but is compromised by two vowel types, viz. /æ/ and /ɑ/. These vowels are long in the perceptual representation of the native AE listeners but substantially shorter, although not as short as any of the lax vowels, in the perceptual representation entertained by the EFL learners. A tentative explanation for this phenomenon might be that some of the PA EFL learners consider the vowels /æ/ and /ɑ/ to be phonetically short, possibly due to exposure to the pronunciation of these vowels in British English, where /æ/ is always short (as it should be given that it is a lax vowel on distributional grounds) and where AE /ɑ/ is pronounced as short /ɔ/ in *doll*. At the same time, Figure 4.7 shows that duration plays a much more prominent role in the perceptual representation of the AE vowels for the EFL learners than for the native listeners. I observed that the PA learners assumed (much) larger duration differences between the AE vowels than were seen in the responses of the native AE listeners. This exaggeration of the duration differences has been reported earlier in the active production of the AE vowels by Arabic EFL speakers from a wide range of regional Arabic varieties (Munro, 2003; see Chapter 5 for details).

4.4.3. PA EFL learners' confusions in depth

In relation to the third question, PA learners' confusions at the vowel level were also accounted for by vowel quality. The most important deviations of the AE vowel space from the natives' norm in nonnative PA mental conceptions are visualized in Figure 4.9. These confusions are argued to require serious attention at the pedagogical level, as they will negatively impact the learner's perception and word recognition skills in EFL and most likely lead to pronunciation errors (see Chapter 5). Among these confusions in perceptual representation, seven confusions (indicated with thick red arrows in Figure 4.9) occurred in more than 20% of the cases (or more). These prominent confusions, however, never involved the tense-lax contrast; rather, the most serious confusions are related to differences in quality within either lax or tense pairs of spectrally adjacent vowels. Learners seem not to discriminate between them correctly and apparently consider them to be very similar to each other. Serious systematic confusions in tense vowel pairs were observed only occasionally (in only 3 out of 24 arrows): /æ/→/e/, /o/→/ɑ/, /ɑ/→/u/, and among mid and low-back vowels only, and even fewer within lax pairs (8.3%).

Curiously, the perceptual representation as entertained by the PA EFL learners is superior even to that held by the native listeners for vowel types in the low back part of the AE vowel space. The results show that the contrast between the vowels /ɑ/ and /ɔ/ is upheld better by the EFL learners, as indicated by the negative numbers (and green arrows) in Figure 4.9. I will ignore this phenomenon in light of the consideration that the poor separation between /ɑ/ and /ɔ/ (and hence the confusion with adjacent vowel types) is caused by the low back vowel merger in Californian English, which is the regional variety of the native AE listeners.

4.4.4. Comparing assimilation to L1 with mapping L2

PA learners show good discrimination of AE vowels in the PAM test when they assimilated them to their L1 vowel inventory and only with difficulty in discriminating adjacent vowels that are not part of the tense-lax contrast (i.e., pairs of spectrally adjacent tense-tense or lax-lax pairs) and do not exist in learners L1. Based on the problematic contrasts in the perceptual assimilation test in Chapter 3, members of these contrasts should yield poor vowel mapping (i.e., larger confusions) in the results of the present chapter. Comparing the results in Figure 3.2 with the results in Figure 4.9 reveals a similar trend. Members of the contrasting pairs in PAM that do not have a (near)counterpart in L1 yielded serious mapping misconceptions but with different AE vowels (than those of the contrasting pairs in PAM), mainly with vowels that contrast in the tense/lax feature. In other words, none of the confusing pairs in the PAM test were among the serious confusions in mapping AE vowels according to the results of the present chapter, and misconceptions only were in adjacent vowels that contrast in the tense/lax feature with vowels lacking in L1. This calls attention to two things. First, there is the role of duration in organizing the perceptual mapping in the mind of the nonnative learners in comparison to the native norm. The PA results show a significant correlation between the length of the stimuli and the tenseness feature of the chosen vowels. Second, there is the role of L1 interference in warping the perceptual mapping of EFL. Here, the results show that most of the confusions involve (at least) one L2 vowel that has no clear counterpart in L1 or confusions between two vowels that do not exist in L1, which in return suggest a clear L1 interference in filtering L2 features. These L2 vowels with no counterpart are considered the most marked vowels (MDH term) in L1 and were automatically the most difficult to discriminate, yielding serious confusions. Less marked or unmarked vowels either yielded less serious confusions or no confusion at all.

The good mapping of the high tense and lax vowels for the nonnative learners in the present study conforms with other studies that reported Arabic vowel inventory to have such features, and that differences in Arabic vowel inventory are not exclusive to duration but also involve differences in vowel quality for the short/long counterparts. A detailed comparison of the native results with the nonnative results shows that native listeners attach more equal weight to quality and duration in the mental representation of the AE vowels, whereas nonnatives depend more on duration as a reliable cue for mapping AE vowels, while only partially taking quality into consideration. Arguably, even if incorrect representations based on duration differences comprise most of the serious confusions (in the tense-lax contrast), they can still be corrected more easily than confusions in quality differences if targeted and explicit learning tasks in the teaching process are employed.

Despite the fact that AE low back vowels are supposed to be NEW sounds for PA learners by definition, they were more correctly mapped by nonnative learners than by the natives. However, this good mapping will not help L2 learners much because many American native speakers have merged these vowels into a single category. On the other hand, other NEW sounds were the most problematic to discriminate and the most difficult to be correctly mapped, which support L2LP postulations in this context for NEW sounds.

4.5. Conclusion

According to the PA results for this task, the Palestinian Arabic learners, in their mental perception, possess a wrong conception of the AE vowel system, with crude color contrast and bisection of stimuli by length. The tense/lax contrasts are not preserved, with the partial exception of the high vowel region. For the PA learners, the location of the centroids of AE tense /i/ and /u/ is clearly more peripheral than the more centralized centroids that are observed for the lax counterparts /ɪ/ and /ʊ/. This is an important insight. I interpret this finding as an indication that the contrast between the short and long point vowels in (Palestinian) Arabic is not simply a matter of vowel length. The results suggest that the contrast is carried not only by vowel duration but also by vowel quality, such that the short counterparts are also lowered and centralized. This would be tantamount to saying that Palestinian Arabic has a tense-lax contrast in its vowel inventory rather than just a length contrast. Similar claims have been made for other varieties of Arabic (Al-Ani, 1970; Watson, 2002). The consequence of this finding would be that the English tense-lax contrast should not present a special problem for Palestinian Arabic learners of English, at least in so far as the high vowels are concerned.

The PA learners have a fairly correct representation of the tense high vowels /i/ and /u/. These tense vowels are properly separated from their lax counterparts /ɪ/ and /ʊ/, respectively. However, then the problems begin. The lax high-mid vowels take up far too much space and usurp the area that should be allotted to the mid vowels /ɛ/ and /ɔ/. Presumably, the PA learners believe that the vowel in *pot* is a short/lax mid vowel /ɔ/ as in British English (rather than the open tense vowel /ɑ/ of AE). As a result of this misconception, the entire mid section of the vowel space is taken up by the conjoined categories /ɪ/ & /ɛ/ for short front vowels and /ʊ/ & /ɑ/ for short back vowels. Long mid vowels are identified as tokens of AE tense /e/ in the front half, and as tense /o/ in the back half of the mid section. Then, all low vowels, including /æ/, together form one conjoined vowel category /ʌ/ (which is not even fully open) when the duration is short. Low vowels are identified as /æ/ only if extremely front and long.

Although all AE vowels were used as response alternatives by the PA participants and no missing islands appeared (as noted in Figure 4.6), the perceptual assimilation of AE pure vowels is clearly dissected to mirror a tense-lax contrast influenced by the short-long contrasts that exist in PA, and the L2 short/lax mid vowels section, which has no (near)counterparts in PA, are perceived as extensions to their nearest short/lax high vowels. Mainly, learners were only sensitive to vowel duration rather than vowel quality. This in return indicates a massive uncertainty, in the mind of PA learners, of what the AE vowels should sound-like. The present results provide clear evidence of L1 interference in L2 perceptual assimilation and identification. Moreover, some studies have reported that the listeners' perception of native and nonnative sounds acoustically match the properties of their native language (Escudero, 2005; Escudero & Vasiliev, 2011; Escudero & Williams, 2011). Nonetheless, there are some pieces of evidence that suggest a developing language faculty in the minds of the PA learners where they clearly differentiate between adjacent vowels; their configuration of the slight differences between these vowels will eventually make them able to adjust their perceptual representation of the AE vowels correctly, especially for the tense/lax contrast.

The results from the previous chapter also support these tendencies. In the previous experiment (PAM-test), I reported that most of the difficulties in perceiving the AE vowel inventory were in vowels that do not have a clear counterpart in PA rather than in vowels with (near)equivalents. It is clear that duration is depended upon as the primary cue that keeps the tense and lax vowels apart for the PA learners, rather than vowel color, which is what natives do depend on. Nevertheless, the PAM results indicate that some reallocations of stimuli to their (correct) places involve sensitivity to vowel color, not just to duration, for the nonnative PA learners.

In the broader picture, these results should have an echo in the production part of the L2. As a rule of thumb, learners could not possibly acquire a good production of the AE vowels unless their perceptual representation of the L2 vowel phonemes is correct, which principle should be applicable to any L2. The results from this chapter and the previous one provide a view on the mental image that L2 learners have of the vowels in the TL, which is helpful in confirming the consistency and the correlation of making errors in both perception and production. Accordingly, I expect substantial errors in speech production for the PA learners, which can be connected back to their perception mistakes. Presumably, such production mistakes can be fixed if a better perception is developed.

CHAPTER FIVE

VOWEL PRODUCTION OF THE AE MONOPHTHONGS BY PALESTINIAN-ARABIC LEARNERS OF ENGLISH AS A FOREIGN LANGUAGE

5.1. Introduction

This chapter is organized as follows: Section 5.1 provides a general description of the study that contextualizes it in relation to the available literature. This section also presents my questions and hypotheses related to the vowel production task. In Section 5.2, the employed method is described as implemented in this production task by highlighting the procedures and the actual steps of the test. Next, in Section 5.3, the results of the test are presented for the native and the nonnative AE speakers. Lastly, Section 5.4 discusses the implications of this task and compares the results of the two groups of participants.

5.1.1. Overview

One of the primary goals for second language (L2) students is to become proficient speakers, or at the very least to be able to communicate intelligibly in the L2. Researchers affirm that pronunciation is among the most important skills in L2 acquisition. Numerous studies (e.g., Cenoz & Lecumberri, 1999; Gilakjani, 2011; Munro, 2010) have confirmed the importance of this skill as a complex phenomenon in which many factors play a role. Among the factors that affect the acquisition of accurate or proficient L2 pronunciation are individual differences (e.g., auditory acuity, working memory), motivational differences, age of acquisition, exposure to L2, innate phonetic ability, identity, etc. More specifically, the type and amount of input are important for the pronunciation skill (e.g., Gass, 1997; Mackey, 1999). Learners who are exposed to English as L2 in a school setting receive audible input either from recorded native speakers or, most of the time, the main source of input is typically the nonnative teachers who have accented pronunciation. Therefore, not only does the learners' L1 affect the L2 production of sounds, but the L2 input source may also not be native. The nonnative EFL teacher functions as a role model, which may result in a compromised perception and production of the target sounds on the part of the EFL learners. Teachers can help learners predict and grasp the differences across the competing sound inventories by comparing important and problematic elements in L2 AE sounds to their counterparts (if they exist) in L1 or by explaining the nature of the *new* sounds that need to be added to the learners' vowel space, either as newly recognized areas in the vowel space or by splitting an already existing category through some learning tasks.

Within the SLM framework, studies (e.g., Rauber et al., 2005) have reported that L1 and L2 sounds overlap as they occur in the *same phonological space* (a term that was changed in the latest version of SLM-r to "common phonetic space", see Flege & Bohn, 2021: Chapter 1) and have near or close phonological attributes and measurements. Accordingly, L1 phonetic categories (being the first to have been acquired) will influence learners' abilities to categorize L2 sounds as *new, similar,* or *identical* and interfere by filtering or warping their perception and production with moderate, most, or least difficulty, respectively. The L2LP model, however, predicts that PA learners would find the NEW sounds the most difficult to acquire, while SIMILAR sounds should be the least difficult (see Section 2.3.4).

Traditionally, Arabic long and short vowels were believed to be at the extreme positions of the vowel triangle and to differ only in duration (Al-Ani, 1970; Holes, 2004; Mustafawi, 2018). However, numerous studies have reported that Arabic long vowels differ both qualitatively and quantitatively from their short counterparts, e.g., Watson (2002) on Egyptian Arabic; Fathi & Qassim (2020) on Iraqi Arabic; Saadah (2011) on Palestinian Arabic; Almbark (2012) on Syrian Arabic, and Kalaldeh (2018) on Jordanian Arabic, among others. Arabic has only 6 vowels, which will overlap with the 11 monophthongal vowels in the American English inventory. As a result, the L2 learners' pronunciation errors can be viewed as reflections of the sound inventories' interaction (interlanguage) within the learner's mind, which is affected not only by the quantity and duration of their L1 vowel inventory but also by other qualitative features exclusive to different L1 Arabic varieties. In light of this assumption, I present the purpose of this study next.

5.1.2. Background of the study

To the best of my knowledge, no earlier study has acoustically tested the PA variety production of English sounds in general and of AE in particular. Meanwhile, studies on the production of other Arabic L1 varieties of English are few and far between. Therefore, this chapter (study 3) tests the PA production of the AE monophthongs. It has been decided so in order to maintain the homogeneity of comparison between the vowel perceptual assimilation and mapping tests (studies 1 & 2) and this vowel production test. In this chapter, I test the 11 AE monophthongs as produced by L1 PA learners via a production task and compare them to learners' results available for native English (Wang & Van Heuven, 2006).

In the next chapter (Chapter 6), I will also compare the results of this chapter with those found in the literature for EFL learners with Arabic L1 backgrounds other than PA (Alqarni, 2018; Khalil, 2014; Koffi, 2021; Munro, 1993).[34] Perhaps the most common and shared results among the studies that approached the production of English vowels by Arabic as L1 speakers are that they struggle with AE back vowels and tend to produce more centralized and shorter vowels than the natives do. Among the general findings, there is insufficient contrast between adjacent pairs of English high-, mid-, and low-front vowels produced by different Arabic L1 speakers (Munro, 1993).

Several studies have been conducted to examine the acoustic characteristics of native (Hillenbrand et al., 1995; Peterson & Barney, 1952) versus foreign-accented productions of English sounds (e.g., Hagiwara, 1997; Lindblom, 1990). Many of these studies have been conducted on this topic with different L1 learners of EFL, but the number of studies on EFL with Arabic as L1 is still lacking and does not cover all Arabic L1 varieties. Some of these studies approached the difficulties that Arabic L1 learners of EFL face both in perception and production (e.g., Almbark, 2012; Evans & Alshangiti, 2018; Nikolova-Simic, 2010), while other studies addressed EFL production only (e.g., Alqirni, 2018; Brown and Oyer, 2013; Hubais and Pillai, 2010; Khalil, 2014; Koffi, 2021; Munro, 1993, among others).

These studies relied on vowel inventories of participants with different L1 Arabic varieties and their production of English sounds as governed by these inventories. Therefore, in this study, I will check the effect of other L1 Arabic varieties on the production of AE vowels that have not been investigated before, namely, Palestinian Arabic. The vowel inventories for different Arabic varieties differ among them and, of course, differ from the English vowel inventory in different features. Therefore, the possibility exists that different L1 Arabic varieties will yield different effects on L2 vowel production. I will start by summarizing the findings of some of these studies below.

In Munro (1993), ten English vowels were examined in /bVt/ and /bVd/ contexts by two adult groups composed of 23 native AE speakers and 23 Arabic L1 late learners of EFL. The EFL learners were from different Arab countries: Kuwaiti (6), Jordanian (S), Sudanese (4),

---

[34] Munro (1993) included two PA persons in a set of 23 L1 Arabic EFL speakers. However, no speaker-individual measurements were given.

Saudi Arabian (3), Syrian (2), Palestinian (2), and Egyptian (1) who stayed in the USA for several years. According to an acoustic analysis of F1, F2, and duration, the study reports that the two groups differ greatly in at least one parameter for almost every vowel and that the nonnatives produced English vowels that are strongly shifted toward their Arabic counterparts. The results showed that accented vowels had different patterns of formant movement, indicating that most Arabic speakers had not learned to produce English vowels with native-like patterns of diphthongization. Due to this finding being observed in all L1 Arabic participants, it is likely that learners are not very perceptually sensitive to formant movements or that they notice the movement but cannot produce it. Low correlation has been reported between the length of exposure and the test scores but the study does not give the actual numbers. In addition, a subset of two (out of 23) EFL speakers had F1 and F2 values comparable to native English speakers when producing the five front vowels /i, ɪ, e, ɛ, æ/ that were rated at 75%; the threshold for nativeness (even though the two natives had means of 86 and 92%). Additionally, among the results, Arabic L1 speakers tended to exaggerate the duration differences between English vowels, suggesting that the durational distinctions for English vowel contrasts were based on the Arabic trend to distinguish vowel contrasts based on duration, yet all AE vowels produced by Arabic L1 speakers were consistently shorter than those produced by native American speakers.

Nikolova-Simic (2010) investigated how L1 interference affects the acquisition of vowels by EFL students from Saudi Arabia by comparing Arabic and English vowel inventories. Learners found several challenges, particularly when addressing speech segments that are comparable in Arabic and English, as well as speech segments that do not exist in Arabic but are widely used in English. The study examined only 10 vowels in American English, and no evidence has been reported on whether experience has a positive impact on the perception and production of English vowels by Saudi EFL learners. Among the study results, it reports that the /ɛ/ and /ɔ/ vowels caused the most problems for both groups of beginners and advanced learners in terms of perception, while the vowels /o/ and /ɛ/ were the most problematic for production. Nikolova-Simic (2010) reports that students were more likely to be able to produce and perceive vowels that were common in both languages, such as [i], [u], and [æ]. Vowels that did not differ by more than one feature from their Arabic counterpart were the next easiest to perceive (but not produce), such as [ɪ] and [ʊ]. The most intriguing result of Nikolova-Simic (2010) was that the sounds that are considered allophones in Arabic were difficult to perceive but easy to produce in English. Among the most challenging vowels to perceive, but not produce, are those that are not present in Arabic, such as [ʌ], [ɛ], [e], and [ɔ].

In a production task for 10 nonnative Omani L1 participants, Hubais and Pillai (2010) compared the Omani-accented English vowel space with the native British English vowel space. According to this study, Omani-accented English occupied a vowel space similar to that of British-accented English, but with varying qualities among the participants. It was found, however, that the Omani-accented vowel space was smaller than the native BE vowel space, with different categorizations of English vowels. In Omani production, for instance, the /ɒ/ was found to be higher than /ɔː/ and closer to /uː/. In general, /iː, æ, ɑː, ʌ/ and, in particular, /ɪ, ɜː, ʊ, ɔː/ were produced more fronted than in BE. The fronting of /ɜː/ is said to be attributed to the realization of the following [r] by the speakers of the targeted words. Additionally, /e/ appeared to be shifting toward /ɪ/. As a consequence, Omani-Arabic L1 students of EFL may have incorporated features of Arabic vowels into their production of matching English vowels.

Brown and Oyer (2013) also addressed how L1 Arabic speakers pronounce AE vowels. The study reported on the English vowels as produced by a single L1 Saudi Arabic speaker who was compared with the GAE results in Peterson & Barney (1952). The Arabic speaker of AE produced more centralized F1-F2 formants compared to the GAE natives. The participant's high vowels were lowered, and the back vowels were more centralized, yielding some major vowel clusters, i.e., /e, i/, /ɪ, ɛ/, /o, ʌ/, /ʊ, u/ and /æ, ɑ, ɔ/ that had formant values very close to each other which were also near to other GAE vowels produced by the natives. The results were further analyzed based on F1 and F2 differences separately between the nonnative and the native GAE participants. When comparing the SA production results with GAE natives, the study stresses that most intelligibility issues for the SA speakers would occur between the members of a number of vowel clusters such that SA production of /ɪ, e/ is so close to native GAE production of /ɛ/, /i, e/ to GAE's production of /ɪ/, and /u/ to GAE's /o/ based on the F1 differences between the Saudi participant results and the GAE results.

In another study, Khalil (2014) provided a cross-language comparative study that addressed Egyptian English and compared it to General American English (GAE) (those in Peterson & Barney, 1952). The primary goal of her study is to compare the intelligibility of Egyptian and GAE speakers of English by testing the production of 11 GAE vowels in /hVd/ contexts produced by 10 Egyptian (5 female) participants. The study results reported a high influence of the Egyptian Arabic vowel system on the produced GAE vowels. The difficulty of GAE vowels for Egyptian speakers varied. Some vowels were found easy to pronounce (i.e., /i, u, ʌ/), others were more problematic: /ɪ, ʊ/, and other, mainly low, vowels such as /æ, ɛ, o, ɔ, ɑ/ were the most difficult to pronounce.

5.1.3. Study questions

This chapter aims to assess PA learners' production of AE monophthongs through a qualitative and quantitative analysis by comparing it to native AE speakers' production of vowels to provide a comprehensive description of AE vowels produced by PA learners. Based on the results of the previous two studies (Chapters 3 and 4), many AE vowels were predicted to be problematic for PA learners in production.

With these points in mind, the production task in the present chapter attempted to answer the following research questions:

Q1 In terms of acoustic measurements of formants (quality) and durations (quantity), how do Palestinian Arabic speakers produce the American English vowels compared to (i) native AE speakers?

Q2 Which AE monophthongs are difficult to produce by PA EFL learners?

The main goal of the present chapter is to determine whether the PA production of AE vowels differs from the native norm, and if so, how. Two follow-up questions will be addressed separately in the next chapter (Chapter 6), namely:

Q3 Is there a connection between PA learners' perception errors and their production (how does their perception correlate with their production of AE monophthongs)? and

Q4 Do the PA production results of L2 AE differ from other nonnative Arabic L1 varieties other than PA with regard to the first two formants (F1, F2) and duration of L2 AE?

5.1.4. Hypotheses

The number of contrastive vowels in any two languages being compared may differ (Maddieson, 2013b). As discussed in Chapter 2, PA learners of AE contrast 6 vowels in their L1. Hypothetically speaking then, at least 5 AE monophthongs should be problematic for them. This is expected since Arabic has fewer contrasting vowel phonemes. Additionally, Arabic vowel phonemes have a wider range available for allophonic variation than in, for example, English or French, although this does not automatically mean that the extra space is used.

In Chapters 3 and 4, it was stressed that L2 vowels are perceived phonemically as contrasting vowels or as contextually induced variations (allophones), where learners usually perceive some L2 vowels as instances of L1 or as their closest L2 counterparts. Intuitively, L2 learners will also produce these vowels in the same manner. Moreover, knowing how or whether L1 and L2 vowels are perceptually connected to one another is a key factor in predicting how learners will produce L2 vowels. Therefore, based on the results in chapters 3 & 4, the following hypotheses can be formulated:

H1. PA learners' production of the AE vowel space is incorrect, especially for the mid-back vowels. Less problematic are the high vowels with tense-lax contrasts between /i-ɪ/ and /u-ʊ/, which will still be differentiated based on duration differences rather than a difference in vowel quality.

H2. PA learners' production of the AE vowel space will gravitate toward the constellation of the PA L1 triangle-shaped vowel system.

H2.1. Many of the mid-section AE L2 vowels will be affected.

H2.2. The results of H2.1 (if not falsified) will be that learners will cluster/lump the mid-back vowels together.

H3. There will be various differences between the PA test results and the findings of such tests on other Arabic L1 varieties on the grounds that each variety has its own exclusive features. (This hypothesis will be explicitly tested in Chapter 6.)

In light of these hypotheses, I also have to consider the possibility that based on L2 vowel inventory development and upon the establishment of new L2 vowel categories, the perceived relationship between L1 and L2 could also change and in return may affect vowel production as well. This possibility is made explicit in more recent versions of the Perceptual Assimilation Model (PAM-L2, e.g., Tyler, 2019).

5.2. Method

5.2.1. Preliminaries

As explained before, all tests and experiments reported on in this dissertation were carried out on the same occasion. First, participants filled in the Language Background Questionnaire, then performed the perceptual assimilation task (Chapter 3) and took part in the vowel identification experiment (Chapter 4), and finally recorded the materials of which a small part will be analyzed in the present Chapter 5. Since the participants are the same individuals in all three chapters, I will refer to Chapter 3 for details on the participants. The description of the stimulus materials and the experimental procedures followed during the recording session, however, will be dealt with in detail in the present chapter.

5.2.2. Stimulus material

The stimulus materials of the present experiment aimed to test the participant's pronunciation of the monophthongs of American English. In the future, I want to study also other aspects of PA EFL learners' pronunciation of English, so that a fairly comprehensive set of materials was developed, from which only a small part will be analyzed in the present dissertation. The full set of materials did not only comprise the 11 English monophthongs, but also the diphthongs and r-colored vowels. Further, the materials covered all the consonants of English, in onset and coda position, as well as a large selection of initial and final consonant clusters. Finally, I recorded the PA students' renditions of the Aesop fable *North Wind and Sun* (the most widely used passage of connected speech used in phonetics, see the Illustrations of the International Phonetics Association in the IPA Handbook: Ladefoged, 1999; Thelwall & Sa'adeddin, 1999). The fable was recorded once in English and a second time in the learner's native variety of Arabic – as an indication of the speaker's natural fluency.

Our choice of monophthongs is essentially the same as that made by the widely used course on spoken English developed by Celce-Murcia et al. (2010). Their choice, which was subsequently adopted by Yavaş (2011), follows earlier studies by, e.g., Lehiste and Peterson (1961). A few remarks are in order here.

- Celce-Murcia et al. maintain that American English has 14 vowel phonemes (see also del Rosario-Garita et al., 2019). Eleven of these are either simple vowels, transcribed with a single vowel symbol. Eleven of these sounds are either simple phonemes, which means that are not accompanied by a glide movement (/ɪ, ɛ, æ, ɑ, ɔ, ʊ, ʌ/, or vowels with an *adjacent* glide, that is, accompanied by /y/ or /w/: /iy, ey, ow, uw/). The remaining three phonemes are diphthongs formed by a vowel sound followed by a *nonadjacent* glide within the syllable (/ay, aw, ɔy/). When the glide element is spectrally adjacent to the base vowel, the vowel is still considered to be a monophthong, on the assumption that such small offglides may be noticeable but do not cause perceptual confusion with other vowels in the American English vowel inventory. Diphthongs, on the other hand, are characterized by a larger change in vowel quality between onset and offset, such that an intermediate vowel phoneme can be identified between onset and offset of the diphthong, which are therefore called non-adjacent. For instance, the diphthong /ai/ executes a spectral change from open /a/ to close /i/, and traverses the area in the vowel quality space that is taken up by /æ, ɛ, ɪ/. If the glide element is eliminated from a diphthong, it will no longer be distinct from at least one other vowel in the inventory.

For instance, /ai, au/ will become identical to one another, and to the monophthong /ɑ/, and /ɔy/ will coincide with /ɔ/.

- Within the set of 11 monophthongs, a division is made between seven long vowels (/i, e, æ, ɑ, ɔ, o, u/ and four short vowels /ɪ, ɛ, ʊ, ʌ/ (e.g., Lehman & Heffner, 1940; Peterson & Barney, 1952; House, 1961; Wang & Van Heuven, 2006; Celce-Murcia et al., 2010). The short vowel typically assumes articulatory positions close to the center of the vowel space than the long vowels, which assume more peripheral positions, i.e., located on the outer perimeter of the vowel quality space (see also Figure 2.2.A). On the strength of the argument that more muscle activity is needed to move the articulators to the extreme positions than to positions closer to the center of the space, the long vowels have also been called 'tense' vowels, and the short vowels the 'lax' counterparts (e.g., House, 1961). Weak support for the extra muscle activity in long vowels has been reported by Raphael and Bell-Berti (1975). In this dissertation, I will use the terms short-long and lax-tense pairs indiscriminately, to indicate that there is a systematic difference between the members of the pairs in terms of vowel duration as well as peripherality.

The 11 monophthongs were embedded in a fixed /hVd/ consonant frame (*heed* /hid/, *hid* /hɪd/, *hayed* /hed/, *head* /hɛd/, *had* /æ/, …) following established practice developed by e.g., Peterson and Barney (1952). Because some of the /hVd/ items yield rather unfamiliar words or names, everyday words that rhyme with the /hVd/-item (*need*, *kid*, *played*, *bed*, *bad*, …) were also included in the materials. All words were embedded in medial position a fixed carrier *Now say … again*, i.e., preceded and followed by a vowel sound and with sentence stress on the target word. The full set of vowel stimuli presented is included in Appendix 5.1.

5.2.3. Participants

The PA speakers who participated in this production task were the same 40 participants who performed the previous two tasks (Appendix 3.2). No native control speakers were recorded as part of the present dissertation. Instead, we were given the materials of an earlier study by Wang (2007) and Wang and Van Heuven (2006), which had used the same /hVd/ items in the same fixed carrier phrase. Their speakers were 20 native AE university students (10 females,10 males), who were recorded in The Netherlands shortly after their arrival to the country. These native speakers hailed from a great variety of states in the USA. Their data were used as a baseline against the PA participants' results in this production task.

5.2.4. Procedure

Participants were recorded in individual sessions in a quiet room in their school building. The participant was issued a sheet of paper with the structures to be read in large print. They were admonished to take their time, study the list of materials at leisure, and then read each item at a conversational pace, and breathe in after reading each sentence. Items were read in pairs, first the version with the everyday key word, then the /hVd/ item. It was explained to the participant that the /hVd/ item should rhyme with the keyword. All instructions were spoken by the author using the participants' native language (Palestinian Arabic). One item was repeated at the end of the list to avoid list effects. The recordings were made at a sampling rate of 44,100 Hz (16-bit amplitude resolution) using a PC151 Sennheiser headset with an adjustable close-talking microphone and a mounted pop-filter.

5.2.5. Acoustic analysis

The three recordings for all the vowels in Table 5.1 were extracted for each speaker, but only the 11 monophthongs were acoustically analyzed in this thesis to preserve homogeneity between the perception and the production results, as mentioned earlier. As a result, I obtained 40 participants × 11 vowels × 3 repetitions = 1320 vowel tokens to be analyzed.[35] The processes of segmentation and file concatenation were semiautomated with the help of a Praat script (Appendix 5.2). The automated segmentation was exclusively limited to setting boundaries between speech and silence segments, which were subsequently confirmed manually in the spectrograms for each stimulus word within a phrase. The acoustic landmarks under investigation were designated as time intervals using the TextGrid files. A new tier was added to delineate the exact vowel duration interval manually based on waveforms and wide-band spectrograms as well as auditory verification. Afterwards, a fourth tier was added to prepare for the automatic extraction of a maximum of five formants (always including F1 & F2) and duration using a Praat script.[36] This tier contained the number of formants and the cutoff of the frequency band within which the formants should be located. The onset vowel boundary was manually set at the end of the /h/ sound and the beginning of the first periodic waveform, while the end boundary was set at the end of the periodic waveform and the start of the /d/ closure period (see Figure 5.1). To illustrate, the five-digit number in Figure 5.1 (separated with one space) in the fourth tier was added as an instruction to the Praat extraction script used in formant

[35] This count excludes the token of *head*, which was repeated to prevent list-final effects.
[36] The script (see Appendix 5.3) was written by Jos J. A. Pacilly, software engineer at the Leiden University Centre for Linguistics, The Netherlands, whose assistance is gratefully acknowledged.

extraction, where the first number to the left indicates the number of formants that the script should extract within a frequency band ranging from 0 to the upper limit specified by the four-digit number to the right of the segment. Therefore, the '3 3300' segment in Figure 5.1 means that three formants should be found in a frequency band ranging from 0 to 3300. Formant frequencies are extracted by the LPC Burg algorithm in Praat, using the number of formants and frequency band as specified in the 4th tier in the TextGrid; the default values for the other parameters were kept intact: Time step is automatic, Window length = 25 ms, Pre-emphasis from 50 Hz. To prepare the formants for extraction, the formant tracks were visually verified over the spectrograms. Formants were extracted every 10 ms of the vowel duration. The formant values were output by the script as the median of formant frequencies during the measured interval for the entire vowel segment (see script in Appendix 5.3).

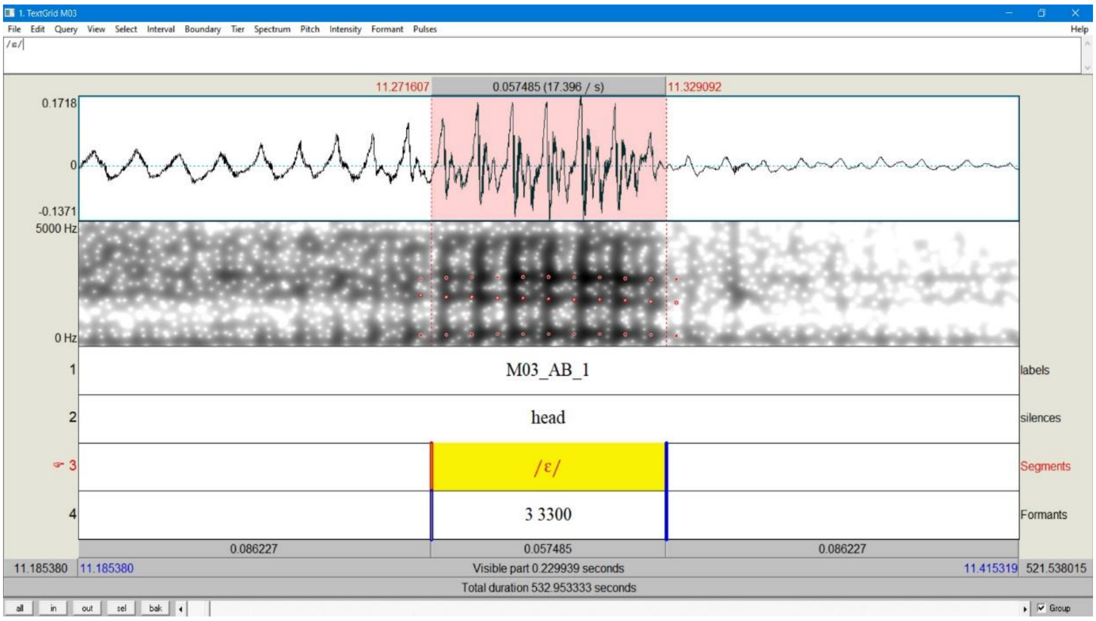

Figure 5.1. Vowel boundaries manually set for the vowel /ɛ/ spoken by a male Palestinian Arabic learner of English as a foreign language. Three formants will be extracted in a 0 to 3300 Hz frequency band.

By default, I tried to find five formants between 0 and 5000 Hz for male speakers and 5500 Hz for females. When no satisfactory overlap of formant tracks over the formant bands in the spectrogram was obtained, I changed the model order of the LPC algorithm in a trial-and-error process by reducing the number of formants (down to three or even two) while lowering the cutoff of the frequency band until a satisfactory visual match was obtained.

The formant (Hz) and duration (ms) values were extracted and saved for further offline statistical analysis. Outliers were spotted in subsequent data analysis by calculating the formant means and SDs separately for each vowel group and each gender group. Any value at ± 1 SD of the mean value was manually inspected to ensure that no error had occurred before or during the extraction process. The data were further categorized based on the carrier words (everyday word vs. /hV(r)d/ carrier). Formant and duration values are plotted below and were generated using the Visible Vowels facility (Heeringa & Van de Velde, 2018).

5.3. Results

I will first present a rather informal overview of the results obtained from the acoustic measurements performed on the recordings of the 40 Palestinian Arabic learners of English as a foreign language. This will be a presentation of the mean F1, F2, and duration values for the eleven target vowels separately for the 20 male and 20 female EFL learners. It will appear that there are no specific effects of gender on the results, so it makes sense to aggregate the results over all male and female learners together. Then, in the next section (Section 5.3.2), I will present a more formal analysis of the data to back up the informal observations in Section 5.3.1 with inferential statistics. Control data from a group of native speakers of American English (10 males, 10 females), which were made available to us, are briefly presented in Section 5.3.3. Then, in Section 5.3.4, I will compare the learners' results with control data obtained through a multivariate analysis, which will allow me to identify the vowels and the (lack of) contrasts among them that potentially compromise the learners' intelligibility in the target language.

5.3.1. Preliminary observations

The PA vowel production results for the AE monophthongs F1, F2, and duration means are presented in Table 5.1. Further statistics (including their SDs, Ranges, etc.) are available in Appendix 5.4.

Table 5.1. Mean F1, F2 (Hz), and duration (ms) of the eleven pure vowels of American English produced by 40 Palestinian Arabic speakers of English as a foreign language, separated by gender.

		American English Monophthongs										
		i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
Male	F1	411	496	510	507	805	654	607	541	483	451	689
	F2	2664	2230	2436	2170	1796	1496	1403	1384	1598	1411	1579
	Dur	98	57	136	67	114	100	152	113	75	100	68
Female	F1	364	442	453	456	666	530	529	481	411	399	614
	F2	2253	1959	2076	1870	1588	1278	1171	1160	1273	1119	1311
	Dur	93	60	133	68	107	88	133	103	81	95	67

Figure 5.2 below plots the vowel formants for male and female PA participants separately based on the raw data summarized in Table 5.2. Difference between /hVd/ words and everyday words was negligible (for inferential statistics, see Section 5.3.2); therefore, they were all

lumped together below. Additionally, in Figure 5.2, I graphically illustrate the spectral shift between the male and female results. There is a consistent correspondence in terms of the relative positions and overlaps among ellipses for the male and female plots, except, of course, that the female formants have shifted to approximately 1 Bark higher than the males, which is physically predictable because females have smaller vocal tract sizes that produce higher formant values. Such physical differences, as a result, are manifested in altering the formant frequencies between female and male speakers of the same language as they produce different frequencies of the same phonological vowels. According to the source-filter theory of speech (Fant, 1960), the formant frequencies of the same vowel can be up to 17% higher in female sound production than in males. My results show a 15% higher shift for the female formants.

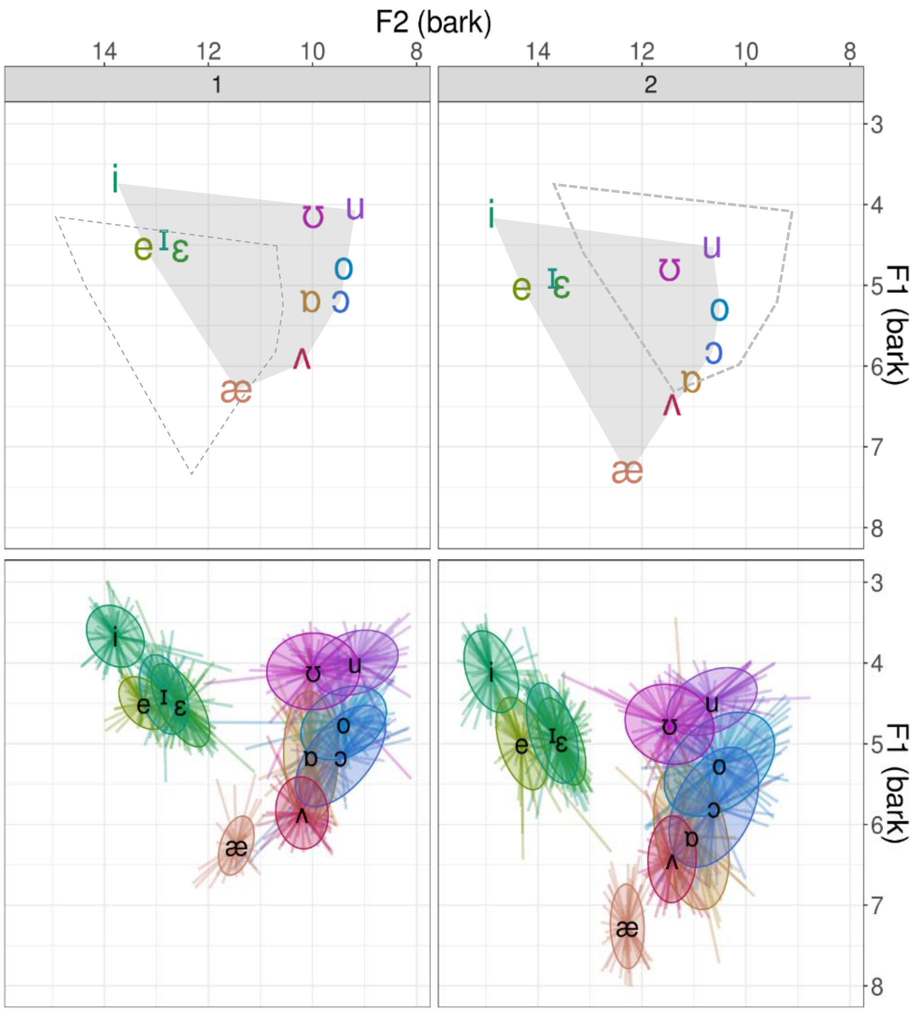

Figure 5.2. Top row: Vowel centroids in vowel space (F1 vertically, F2 horizontally in Bark) for 11 American English monophthongs produced by adolescent Palestinian Arabic EFL learners, left for men, right for women (20 speakers per gender, 1 to 3 tokens per vowel per speaker). The convex hull of the vowel configuration is drawn as a filled polygon. Dashed contours represent the convex hull of the opposite gender. Bottom row: As top row but with all vowel tokens plotted and linked to centroids. Spreading ellipses have been drawn at ± 1 standard deviation along the two principal components of each scatter cloud and include (theoretically) 46% of the most typical data points per vowel type. [Plots produced with Visible Vowels, Heeringa & Van de Velde, 2018].

The overlapping configurations of the adjacent vowels are expected to negatively affect the speech intelligibility for the PA EFL production. It seems that PA learners matched the AE vowels to their nearest L1 vowels. Interestingly, the results show that the vowels that have no counterparts in Arabic have been clustered around their closest match of AE that has a counterpart in Arabic. For example, the AE near-high near-front unrounded lax vowel /ɪ/ that has a near counterpart in Arabic (high front short /i/) has attracted its adjacent vowels toward it, which results in almost an overlap with /e, ɛ/. The PA high front area has only two vowels (/i/and/i:/). It seems that short PA /i/ was realized as the most reasonable spectral substitute for AE /ɪ, e, ɛ/. Surely, duration will keep AE /e/ apart from the other two lax vowels. This phenomenon shows that PA EFL learners are able to split their spectral and temporal properties. Spectrally, PA learners consider these three vowels the same as PA /i/ and have not yet figured out that they should lengthen PA /i/ when pronouncing AE /e/ (the alternative strategy, substituting long PA /i/, is perceived as too deviant by the learners). The same situation is applicable to the /u, ʊ/ cluster but, for these two vowels, the participants can distinguish between them on durational grounds, which is not the case for the cluster in the low-back area. The clustering found for the AE mid- and low-back vowels /o, ɔ, ɑ, ʌ/ suggests that these vowels are totally confused and unfaithfully produced by the PA speakers. These issues will also be discussed in relation to AE vowel durations as produced by PA participants in the following sections.

The mean duration for each AE vowel as produced by PA EFL speakers and the breakdown by gender (1 = male, 2 = female) is shown in Table 5.1 and visualized in Figure 5.3. For more details, see Appendix 5.4.

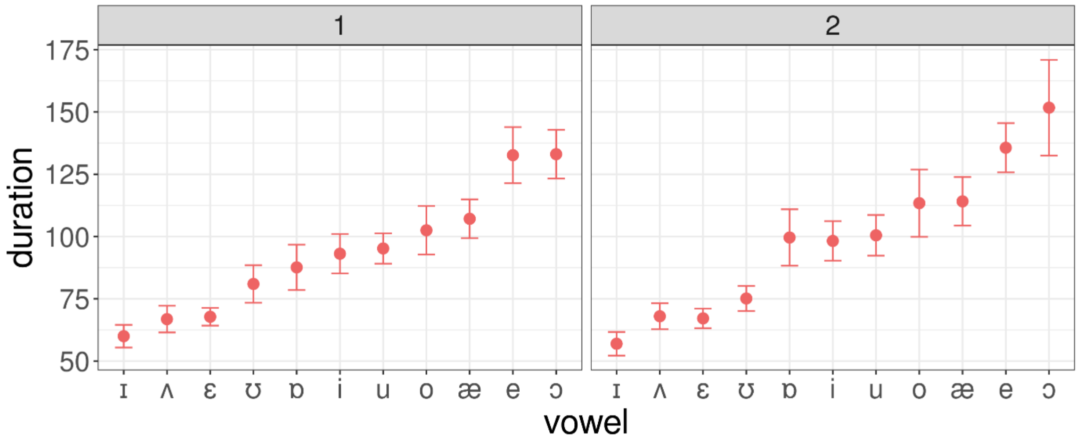

Figure 5.3. Duration (mean and standard error, ms) of 11 American English monophthongs produced by 20 male (left) and 20 female (right) Palestinian Arabic adolescent EFL learners.

No obvious differences in durations are noted between genders, as they show almost the same pattern of vowel durations (for inferential statistics, see the next section). The results in Figure 5.3 show great consistency for the vowel sequence based on duration. Moreover, all PA participants kept the lax/short and tense/long vowels clearly separated, as the lax vowels were produced with obviously shorter durations than tense vowels, as the AE natives would do, especially for the PA female participants.

This can be expected given that Arabic has a short-long contrast in the vowel inventory. Similar findings, however, have been reported in other EFL studies with L1s that do not have a length contrast in their vowel systems (e.g., Perwitasari, 2019 for Sundanese and Javanese learners of American English).

5.3.2. Inferential statistics

For the dataset to be analyzed, I included the token of the keyword, as well as the first and second tokens of the /hVd/word. In principle, the repetition of the word *head* at the end of the list (included to avoid list-final prosody) was excluded, except when the regular token was missing (e.g., due to incorrect pronunciation). In that case, the list-final token was substituted for the regular token. The nominal number of tokens per vowel should therefore be 120, i.e., 40 speakers (20 male, 20 female) × 3 tokens. Eighty tokens per vowel type were recorded in /hVd/ context, 40 tokens were pronounced in /CVd/ keywords. A number of tokens had to be rejected, mainly because the learner mispronounced the vowel in a way that could not be explained as phonological interference from the speaker's native language. The most frequent mispronunciation was found for the /hVd/ item *hawed*, which was realized as /au/ (*how'd*) instead of /ɔ/. The total number of missing or rejected tokens was 92, which amounts to 7.0% of the nominal total number, i.e., 1320. The count of tokens is summarized in Table 5.2.

Table 5.2. Number of vowel tokens suitable for statistical analysis broken down by gender and by vowel type.

Gender	Vowel											Total
	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	
Male	59	59	55	60	58	60	55	49	52	50	56	613
Female	59	55	54	60	60	55	56	53	51	54	58	615
Total	118	114	109	120	118	115	111	102	103	104	114	1228
Missing	2	6	11	0	2	5	9	18	17	16	6	92

Since the design of the experiment is within-participants, the effects of Vowel type and Context on the acoustic variables should be evaluated through a Repeated Measures Analysis of Variance. This technique requires that all cells in the design be filled for all participants. In the PA results, only 8 out of 40 speakers had produced the required full set of 33 (11 vowel types × 3 tokens), which yields too small a sample of speakers for a sufficiently powerful statistical analysis. Empty cells in the data matrix were therefore reduced by aggregating (i.e., averaging over) the nominally two /hVd/ tokens. This decreased the nominal number of datapoints by one-third, i.e., to 440 /hVd/ (6 missing data points, or 1%) and an equal number of /CVd/ points (19 missing, or 4%). The 25 remaining missing data points (= 3%) were restored by noniterative two-dimensional imputation. I computed the marginal means for rows (= speakers) and columns (= vowels), separately for the two contexts (/CVd/ keywords, /hVd/ targets) in the data matrix, skipping the cells with missing values. I then determined per speaker, per vowel, and per context how much the marginals deviated from the grand mean and then replaced the empty cell by the grand mean plus the deviation in the three dimensions. The imputed values were thus equal to what would be predicted by linear addition of the speaker effect, the vowel, and the context effect with no adjustment for possible interactions. Through this imputation, the mean values of speakers, vowels and contexts are the same in the original dataset with missing values and in the restored dataset. In the remaining statistical analyses, the restored dataset will be used. The analyses will be graphically illustrated with interaction diagrams. Full numerical data underlying the breakdowns in the graphs can be found in Appendix 5.4. Table 5.3 summarizes the RM-ANOVAs for F1 frequency (after conversion to Bark), F2 frequency (in Bark), and Vowel duration.

Table 5.3. Summary of RM-ANOVA. The dependent variables are F1, F2, and vowel duration. Within-participant factors are Vowel type and Context (/hVd/, /CVd/). The between-participant factor is Gender of speaker. All main effects and interactions are listed. Nominal degrees of freedom are reported, but *p*-values were computed after Greenhouse-Geiser correction (ε correction coefficient is specified). Significant effects and interactions (α = .050) are in highlighted cells. When the effect size pη^2 ≥ .100, the cell entry is also bolded.

Tests of Within-Subjects Effects		First formant			Second formant			Duration		
Effect/Interaction	$df_{1,2}$	F	p	pη^2	F	p	pη^2	F	p	pη^2
Vowel	10, 380	274.5	**< .001**	**.878**	599.5	**< .001**	**.940**	79.1	**< .001**	**.675**
				ε = .630			ε = .423			ε = .464
Vowel * Gender	10, 380	4.1	< .001	.098	1.9	.103	.048	1.2	.294	.032
				ε = .654			ε = .659			ε = .324
Context	1, 38	.1	.775	.002	17.5	**< .001**	**.315**	29.9	**< .001**	**.440**
Context * Gender	1, 38	.3	.573	.008	.7	.418	.017	13.9	**.001**	**.268**
Vowel * Context	10, 380	8.6	**< .001**	**.185**	9.8	**< .001**	**.205**	3.5	.016	.084
Vowel * Context * Gender	10, 380	1.5	.176	.038	.6	.771	.015	1.3	.275	.033
Tests of Between-Subjects Effect										
Gender	1, 38	75.2	**< .001**	**.664**	200.0	**< .001**	**.840**	.7	.419	.017

Table 5.3 specifies all possible main effects and interactions, separately for the within-speaker and between-speaker terms. In none of the three analyses was the condition of sphericity met, so I used the Greenhouse–Geisser correction of degrees of freedom as a safety

precaution. In the table, however, I list the nominal degrees of freedom (for the sake of clarity); the *p*-values listed were computed after GG-correction. Except for one, all factors in the RM-ANOVA are dichotomies, which require no post-hoc analyses for multiple contrasts. Vowel type, however, has 11 levels, which were tested pairwise by post-hoc t-tests with Bonferroni correction for multiple comparisons. My criterion for significance is $\alpha = .050$; however, to avoid having to factor in small effects, I made the additional requirement that the effect or interaction should have an effect size of partial eta squared ($p\eta^2$) of at least .100. All significant effects/interactions are highlighted in the table; they are additionally bolded if the $p\eta^2$ requirement is met.

First formant (F1). The effect of Vowel is significant. Bonferroni post-hoc tests indicate that the four (half-)close vowels /i, ɪ, u, ʊ/ have the same F1; so do the semi-diphthongs /e, o/, and the pair /ɑ, ʌ/. The F1 of /ɛ/ and /æ/ differ significantly from one another as well as from any other vowel.

Context has no effect on F1 and does not interact with Gender. However, Context interacts with Vowel. The third-order interaction between Vowel, Context, and Gender is insignificant. Female speakers have higher F1 values than males because they have smaller/shorter resonance cavities, so the Gender effect is predictably significant. Additionally, the Vowel-by-Gender interaction is significant but does not reach the $p\eta^2 \geq .100$ criterion, so that this interaction will be ignored. Figure 5.4 plots the F1 center frequency for the 11 vowels broken down by Gender (row panels) and by Context (column panels). The vowels are ordered from left to right in descending order of F2 frequency as they were found in the AE control data (i.e., I used the same axis layout as in Figure 5.5, which plots F2).

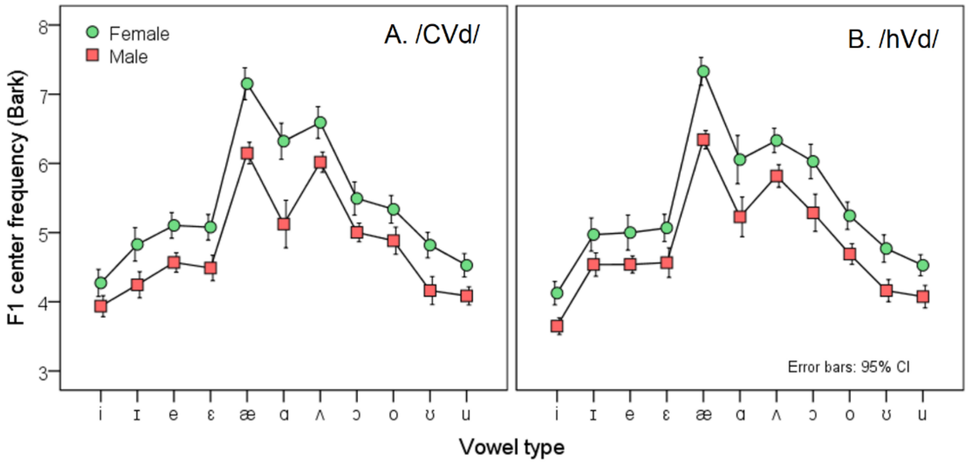

Figure 5.4. Center frequency of F1 (Bark) for 11 American English monophthongs pronounced by 20 male and 20 female Palestinian Arabic learners of English as a foreign language. Error bars are the 95% confidence intervals of the speaker means. Tokens were produced in /CVd/ everyday keywords (panel A) or in /hVd/ context (panel B). Vowels are ordered by descending F2 as determined for a control sample of 20 native speakers of American English.

Post-hoc tests with Bonferroni correction for multiple comparisons (α = .050) show that 4 subsets should be distinguished within which vowels do not differ from each other in terms of F1 (in ascending order of F1): /i, u^{1}, ʊ^{1,2}, ɪ^{1,2,3}, ɛ^{2,3}, e^{2,3}, o^{3}, ɔ^{4}, ɑ^{4}, ʌ, æ/. Vowels with the same superscript number do not differ significantly from one another. Subsets may overlap, e.g., /ɪ/ is a member of three partially overlapping subsets: {u, ʊ, ɪ}, {ʊ, ɪ, ɛ, e}, {ɪ, ɛ, e, o}. The vowels /i, ʌ, æ/ differ significantly from each other and from all other vowels.

The effect of Context is very small visually and can be seen only for the open vowels /æ, ɑ, ʌ/, whose somewhat discrepant F1 in the two contexts would explain the significant Vowel-by-Context interaction.

Second formant (F2). Figure 5.5 plots the F2 center frequency for the 11 vowels broken down by Gender and Context. The vowels are ordered from left to right in descending magnitude of F2 as found in the AE control data.

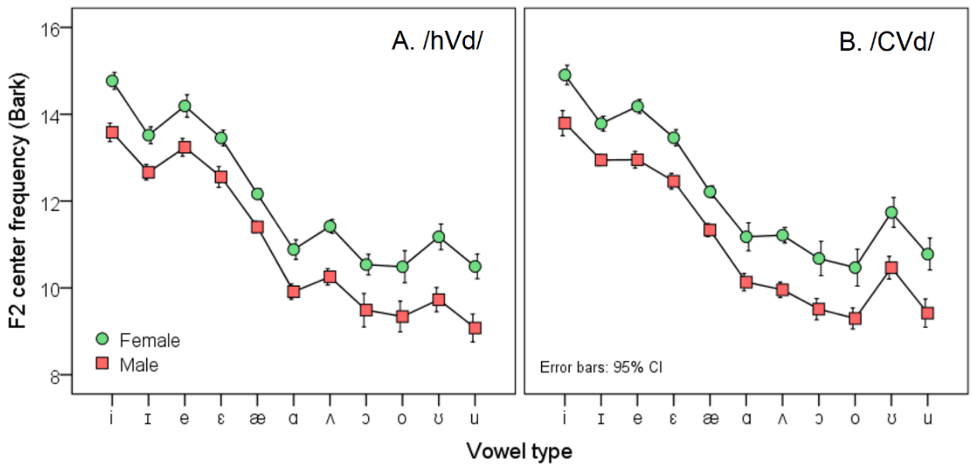

Figure 5.5. Center frequency of F2 (Barks) for 11 English monophthongs pronounced in /hVd/ words and in rhyming everyday keywords (/CVd/) by Palestinian Arabic EFL learners, broken down by Gender and by Context. Error bars include the 95% confidence interval of the speaker means. Vowels are ordered from left to right by descending F2 as found for American control data.

The effect of Vowel is very strong. Post-hoc comparisons show that all vowels differ significantly from each other in terms of F2, with the exception of two nonoverlapping subsets, {ʊ, ʌ, ɑ} and {ɔ, u, o}, within which the members do not differ from one another (listed here in descending order of F2): /i, e, ɪ, ɛ, æ, ʊ^{1}, ʌ^{1}, ɑ^{1}, ɔ^{2}, u^{2}, o^{2}/. Gender is a predictably significant effect (see above). Additionally, the Vowel × Gender interaction reaches significance. Context exerts a small (but significant) effect on F2. Finally, the Vowel × Context interaction reaches significance. No other effects or interactions were found to be significant. Interactions between the main effects (even when significant) are hardly noticeable. There are only two effects that are of importance: the effect of Vowel and the effect of Gender. The Vowel effect is what I am interested in. The effect of Gender will be neutralized in subsequent analyses through z-transformation within speakers (Lobanov normalization). The effect of Context (11.56 Bk for /hVd/ vs 11.66 Bk for /CVd/), even though significant, is indeed very small (± .05 Bark). The automatic classification in the next section will be performed on the /hVd/ tokens only, since this is the only context for which I have native control data (Wang & Van Heuven, 2006). Note that, in contradistinction to the American native control speakers, the PA EFL learners produce a small but significant F2 difference between the vowels /ɑ, ɔ/.

Duration. The effect of Vowel was highly significant. Bonferroni post-hoc pairwise comparisons bear out that in the ascending order of duration /ɪ, ɛ^{1}, ʌ^{1}, ʊ, i^{2}, ɑ^{2}, u^{2}, o^{2,3}, æ^{3}, e^{4}, ɔ^{4}/ there are four subsets within which the vowels do not differ significantly in duration. It should be highlighted that each of the four lax vowels is significantly shorter than any of the phonetically tense vowels. This finding suggests that duration is used by the EFL learners to differentiate lax /ɪ, ɛ, ʌ, ʊ/ from their spectrally nearest (overlapping, see Figure 5.4) tense competitors /e, ɑ, o/.

Vowel durations do not differ significantly between male and female speakers. There is no interaction between Gender and Vowel, nor is there interaction between Gender and any other factor. Target vowels are pronounced longer in everyday keywords (103 ms) than in /hVd/ items (93 ms). Moreover, the Context interacts with the Vowel type. Figure 5.6 shows the interaction. Although there is no second-order interaction between Gender and Vowel duration, there is a remarkable third-order interaction by which the female speakers produce longer vowel durations for the tense vowels (and the males' shorter durations) in the everyday keywords (/CVd/ context) than in the /hVd/ targets (in the latter context, male and female speakers produce equally long durations for each vowel type). This interaction fails to reach significance (see Table 5.3), probably because the second-order interaction between Context and Gender explains the variance, while the male and female tense vowel durations have overlapping confidence intervals even though they differ by some 20 ms. The breakdown of vowel duration by Vowel type, Gender of speaker and Context is shown in Figure 5.6.

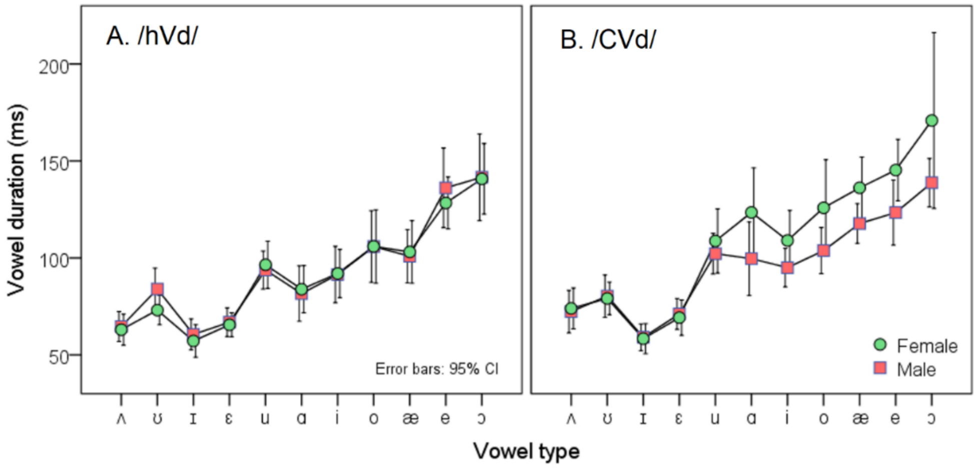

Figure 5.6. Vowel duration (ms) for eleven American English monophthongs produced by male and female Palestinian Arabic learners of English as a foreign language. The vowels are arranged in ascending order of diration as found for 20 American native control speakers. See Figure 5.5 for more information.

5.3.3. American English control data

The control data recorded for native speakers of American English were collected by Wang and Van Heuven (2006) as part of Wang's (2007) doctoral dissertation project. Ten male and ten female adult native speakers of General American English produced the eleven monophthongs of AE in /hVd/ context, in the same carrier phrase that I used for the recording of the Palestinian EFL learners.[37]

In the American native control data, there is no visible spectral overlap between /i/ and /ɪ/, nor between /u/ and /ʊ/. Also, /ɛ/ is clearly separated from /e, ɪ/, and /ʌ/ from the low-back vowel cluster /ɔ, ɑ/. Spectrally, the very substantial overlap between the spreading ellipses for /ɔ, ɑ/ suggest that these two vowels are merged for most of the speakers in the sample. The spectrally overlapping lax vs tense mid vowel pairs /e, ɪ/ and /o, ʊ/ will be distinct by a difference in duration.

[37] I am most grateful to professor Hongyan Wang of Shenzhen University (P.R. China) for making her vowel data available to me. Wang and Van Heuven (2006), and Wang (2007) did not report any results for the tokens of *hawed*, due to the assumption of a complete merger of the low-back vowels in their recordings, but the formant and duration values for the *hawed* tokens were included in the underlying dataset.

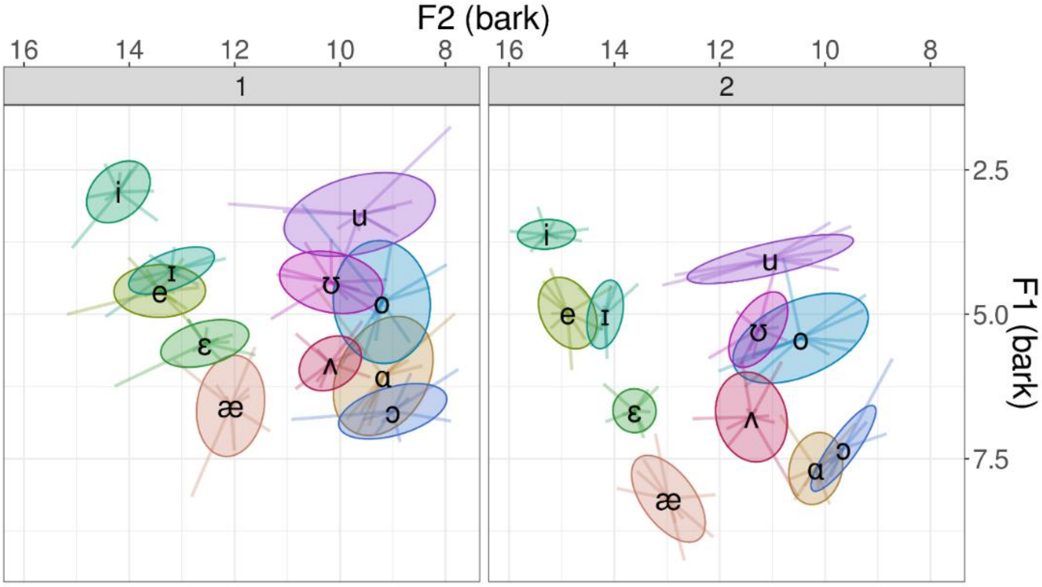

Figure 5.7. American English vowel formant control data. F1 and F2 (Bark) produced in /hVd/ items by 10 men (left panel) and 10 women (right). See Figure 5.2 for details. For information on the recordings see Wang & Van Heuven (2006) and Wang (2007).

Figure 5.7 illustrates the same gender effect on the formant values that was seen in Figure 5.2. It can be observed that the ranges of the formant values are larger for the native control speakers than those found in the EFL learner data (in Figure 5.2). It is unclear at this stage whether the compression of the formant ranges in the EFL speech is due to interference from Arabic (which has no phonological mid vowels) or to a low-level phonetic consequence of shorter vowel durations in nonnative speech. It has been shown in the literature that shorter vowel durations prevent the articulators from reaching their target positions (the target undershoot model), which yields more centralized vowel articulations with formant values gravitating toward those of schwa (e.g., Lindblom, 1963; Koopmans-van Beinum, 1980; Fourakis, 1991; Van Bergem, 1993).

The vowel durations for the monophthongs in the control data are shown in Figure 5.8.

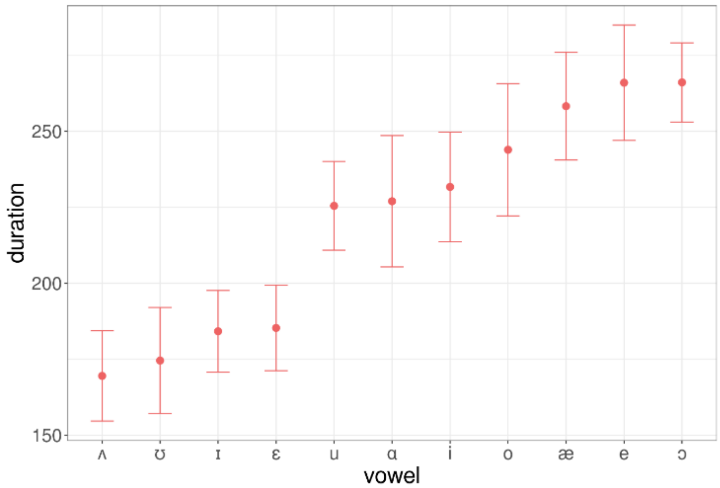

Figure 5.8. American English control vowel durations (in ascending order of duration).

Figure 5.8 shows a clear division between the four lax and short vowels against the seven long and tense vowels. Lax /ʌ/ is the shortest vowel of all. The vowels /æ, ɔ, ɑ/ side with the long vowels in the control data, as can be expected for American English (see section 5.1.2). The native AE mean vowel durations for the /hVd/ context range between 160 and 260 ms, which is more than twice the mean durations realized by the EFL learners in Figure 5.6A. It is unclear, again, whether the EFL learners speak faster overall and therefore produce shorter vowel durations, as well as more reduced formant ranges, or whether their conception of the spectral and durational properties of the AE vowel system is seriously flawed.

In the multivariate analyses in the next section, I will abstract from the absolute differences in the ranges of the formants and durations between L1 and L2 speakers by within-speaker z-normalization. As a result of the normalization, the speaker means and ranges will be rescaled such that each speaker has a mean of zero and a standard deviation of 1 for each of the parameters F1, F2, and vowel duration. However, the relative distances between the vowels will be preserved, and the (smaller) overlap of the spreading ellipses of adjacent vowels will more realistically reflect the linguistically relevant separation of categories.

5.3.4. Multivariate analyses

5.3.4.1. Palestinian Arabic EFL learners

In this section, I will combine the acoustic vowel parameters (F1, F2, and duration) in a multivariate analysis in an attempt to automatically classify the eleven American English vowel types as produced by the PA EFL learners. The RM-ANOVA results above show that the three parameters I measured for the vowel tokens are very sensitive to differences between vowel types. The effect of Vowel in Table 5.3 is strongest for F2 ($p\eta^2 = .940$), second strongest for F1 ($p\eta^2 = .878$) and third strongest for Duration ($p\eta^2 = .675$). All other effects (and interactions) that can be seen in Table 5.3 are small (and often insignificant) in comparison and will not contribute significantly to the automatic classification of the vowels. The one exception is the effect of Gender ($p\eta^2 = .884$ for F2, .640 for F1, insignificant for Duration), but this effect is not linguistically relevant and will be factored out prior to the analysis through Lobanov normalization (z transformation within speakers).

Two different classification algorithms were employed, i.e., Linear Discriminant Analysis (LDA, Klecka, 1981) and Multinomial Logistic Regression Analysis (MLRA, Hosmer & Lemeshow, 1989). This is the same choice that was made in Van Heuven et al. (2020). To test the contribution of vowel duration explicitly, analyses were carried out once with only the spectral parameters F1 and F2 and a second time with duration added. All analyses were performed on the dataset with imputed values for missing cases (see Subsection 5.3.2 above) and subsequent within-speaker z-normalization of the acoustic parameters. Table 5.4 shows the percentage of correctly classified vowel tokens in each of the above conditions. The underlying confusion matrices can be found in Appendix 5.5.

Table 5.4. Percent correct classification by Linear Discriminant Analysis (LDA) and by Multinomial Logistic Regression Analysis (MLRA) of 11 American English vowels produced by Palestinian Arabic learners of English as a foreign language. Percentages are listed for analyses with spectral parameters only (F1, F2) and with spectral parameters plus vowel duration. All predictors were z-normalized within speakers. Columns under Δ specify the difference due to addition of the duration parameter.

LDA			MLRA			Δ MLRA – LDA	
F1, F2	+ Dur	Δ	F1, F2	+ Dur	Δ	F1, F2	+ Dur
63.8	74.7	+10.9	65.8	77.4	+11.6	+2.0	+2.7

On the basis of the two spectral parameters, F1 and F2 formant center frequency, the AE vowels produced by the Palestinian EFL learners can be automatically classified at 64% correct by LDA. Given 11 vowels to choose from, this is approximately 7 times better than chance (= 9% correct). Automatic classification by MLRA is slightly better, with 66% correct. Adding vowel duration as a third predictor improves the correct classification by 11 points for the LDA method and by 12 points for the MLRA method. The classification with duration included as a predictor is better than 8 times chance. It should be noted, however, that these measures merely indicate how well the EFL learners separate their AE vowels. This does not mean that they separate them in the same way American native speakers do. My hypothesis should be that American native speakers keep the vowels more distinct and may well have different centroids and boundaries between adjacent vowel categories.

5.3.4.2. Classifying nonnative vowels by native models

Ideally, all the vowel tokens produced by the PA EFL learners should be offered for perceptual identification. However, since each listener would have to respond to 1228 different stimuli (Table 5.1), even if each token would be presented only once, I decided not to involve human listeners at this stage of the project. Instead, I used the automatic classifiers (see above) as substitutes for a group of human native listeners, as done before by, e.g., Strange et al. (2004), Wang & Van Heuven (2006), and many others. It has been found that an automatic classifier, when properly trained with vowel tokens produced by a representative group of native speakers, divides up the multidimensional vowel space approximately the same way native listeners of the language do. If I then use the classifier to identify new tokens, for instance tokens produced by nonnative speakers, it will identify the new tokens as if they were produced by native speakers. In this way, the LDA classifier is a substitute for the human native listener. In this section, I will use the American native vowel tokens made available to me to train the classifiers and then use the models to identify the vowel tokens produced by the Palestinian Arabic EFL speakers. All predictors were z-normalized within speakers. Leave-one-out cross-classification was applied when the native-speaker model was used to classify the native vowels by LDA.

Table 5.5. Percentage of correct vowel identification by Linear Discriminant Analysis (LDA) and by Multinomial Logistic Regreression Analysis (MLRA), with spectral parameters F1, F2 or with spectral parameters plus vowel duration. All predictors were z-normalized within speakers.

Training set	Test set	LDA			MLRA			Δ MLRA – LDA	
		F1, F2	+ Dur	Δ	F1, F2	+ Dur	Δ	F1, F2	+ Dur
Native speakers	Native speakers	78.6	85.9	+7.3	81.4	90.9	+9.3	+2.8	+5.0
Native speakers	PA EFL learners	55.1	64.0	+8.9	55.8	59.9	+4.1	+ .7	−4.1

As before, all predictors made a significant contribution to the model. With only the two spectral parameters, the native speaker vowels were correctly classified by LDA at 79% correct, which was better than the 8 times chance-level performance (chance = 9% correct) and somewhat better still by MLRA at 81%. Adding (normalized) vowel duration as the third predictor raises the percentage of correct classification to 86 (~9 times better than chance) by LDA and to 91 by MLRA (10 times better than chance). When the native classification models are then applied to the EFL tokens, the performance decreases quite substantially, by LDA to 55% correct for spectral parameters only and to 64% correct when duration is added as a third predictor, and by MLRA to 56% and 60% correct.

If I contrast the correct classification in Table 5.5 (classified by native speaker norms) with the classification when the algorithms were trained on the EFL vowels, in Table 5.4, it can be observed that the scores are considerably better in the latter situation: 64% correct for spectral parameters and 75% correct with duration added for LDA and 66% and 77% correct, respectively, for MLRA. This improvement relative to the results in Table 5.5 can be considered an instance of the Interlanguage Speech Intelligibility Benefit (ISIB). The claim is that listeners who share their native language with the EFL speaker intuitively know the peculiarities imparted by the shared native language to the foreign speech. Listeners who do not share the native language of the EFL speaker cannot use these subtleties. This would also explain why listeners who share the native language with the EFL speaker tend to even outperform native listeners of the target language (Bent & Bradlow, 2003; Wang & Van Heuven, 2015; Van Heuven, 2016).

A graphic summary of the classification results by vowel type is given in Figure 5.9. The figure shows the percentages of correct vowel classification by LDA (panel A) and by MLRA (panel B) when trained on American native vowel tokens and tested on either the same group of native speakers (square markers) and when tested on the PA EFL learners (circles).

Triangular markers represent the condition in which the classifiers were both trained and tested on the PA tokens. Separate curves are drawn for prediction based on spectral parameters only (open markers, dotted lines) and for prediction with vowel duration added (filled markers, solid lines). The vowels are arranged from left to right in ascending order of correct classification (averaged over the two classifiers) obtained in the crucial condition, i.e., where the classifiers are used as a model of the human native listener who responds to the nonnative tokens.

The figure shows that the classification of the vowels /i, ɪ, e, æ/ is at 80% correct or better. Poor classification is observed for the open back vowels /ɑ, ɔ/. This is seen for native and nonnative speakers alike. The remaining vowels are in between the extremes discussed here. With just one exception, it is always the case that vowels are equally well or better classified when duration is included in the predictor set.

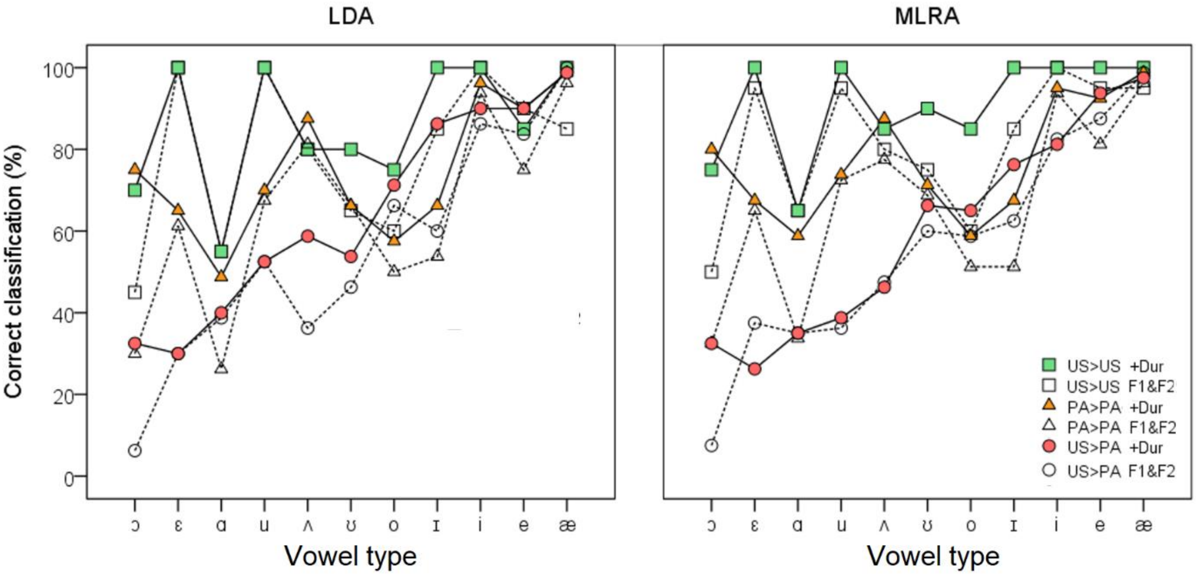

Figure 5.9. Correct classification (%) of 11 American English vowels by Linear Discriminant Analysis (panel A) and by Multinomial Logistic Regression Analysis (panel B) trained and tested on native vowel tokens (20 speakers, square markers, control condition), trained and tested on nonnative tokens produced by 40 Palestinian Arabic learners of EFL (circles), and trained on native tokens and tested on the nonnative tokens (triangles). Classification is performed with spectral parameters only (open symbols) or with vowel duration added (closed symbols). All predictors are z-normalized within speakers. Predictors were z-normalized within speakers.

The largest number of misclassifications (relative to the native classifications) are found for the vowel types /ɔ, ɛ, u/.

A better view of the pronunciation problems of the EFL learners is afforded if the confusion structure in the automatic classification is considered. The full confusion matrices of intended versus actually classified vowels are included in Appendix 5.5. On the basis of the matrices, simplified confusion graphs can be constructed, which highlight the most frequently occurring misclassifications (errors) in the pronunciation of the EFL speakers (which can also be compared with the confusion structure obtained for the native speakers).

Figure 5.10 shows the confusion structure obtained for the classification of the American native-speaker classification based on all three acoustic predictors. The eleven vowels are arranged as in a traditional IPA vowel diagram, with front vowels to the left, back vowels to the right, closed (high) vowels at the top and open (low) vowels at the bottom. The lax and centralized vowels are in the center of the diagram and shaded for visual contrast. The most difficult vowel for the nonnatives is /ɛ/, which is correctly classified in ca. 30% for the nonnatives against (close to) 100% for the natives.

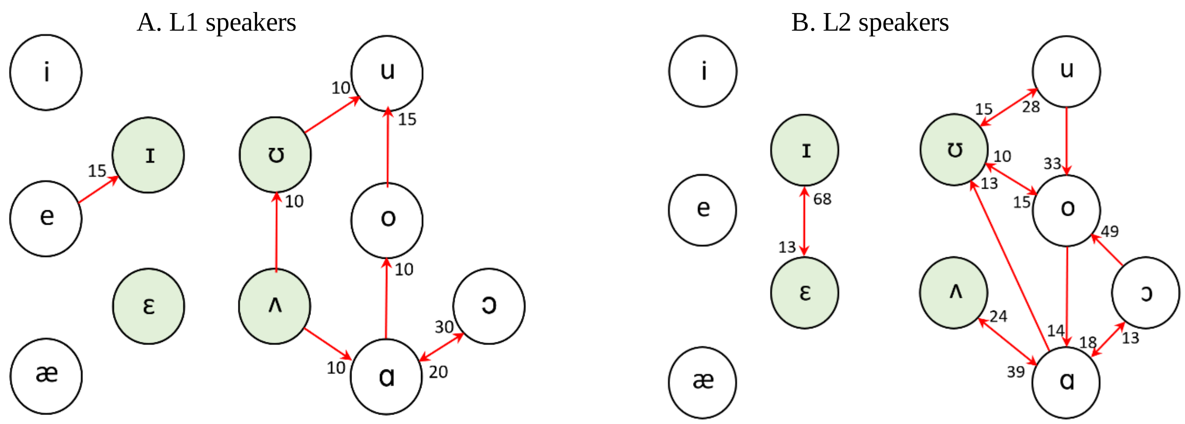

Figure 5.10. Panel A: vowel confusion structure for classification by LDA of eleven American English monophthongs produced by and tested on 20 American native speakers. Predictors were F1, F2 and vowel duration. Confusions < 10% have been omitted. Lax/short vowels in shaded circles. Arrows point away from the intended vowel to the incorrectly identified vowel. The percentage of confusion is indicated at the arrowheads. Panel B presents the same information for the 40 Palestinian Arabic learners of English as a foreign language.

There is relatively little confusion among native-speaker vowels and never more than in 15% of the identifications. The exception here is the substantial confusion between the two low-back vowels /ɑ, ɔ/. Nevertheless, even though these two vowels are implicated in the low-back vowel merger, which is a characteristic of many American dialects, the confusion is relatively mild; classification is much better than chance, $\chi^2(1) = 47.1$ ($p < .001$), and may well benefit from the significant difference in duration between the two vowels (/ɑ/ is ca. 40 ms shorter than /ɔ/ – see Figure 5.8). There is no confusion between nonadjacent vowels or between front and back vowels. However, there is a tendency for back vowels to be more vulnerable to confusion than front vowels.

It is plain that substantial confusion surfaces when the native speaker models are used to classify the vowel tokens produced by the PA EFL learners (Figure 5.10, panel B). However, the confusion is limited to specific clusters of adjacent vowels only. Moreover, there is never any confusion between front and back vowels. The most prominent confusion is between the lax front vowels /ɪ, ɛ/. The asymmetry between these two vowels suggests that the half open /ɛ/ is articulated too closed.

Confusion is much more prevalent between pairs of back vowels. The confusion structure suggests that the lax back vowels are insufficiently differentiated from the adjacent tense back vowels. Typically, half close /ʊ/ is poorly contrasted with /u, o/, while half open /ʌ/ is poorly distinguished from open /ɑ/. The asymmetries between members of confused pairs suggest that /ʌ/ is pronounced too far back, while /u/ should be raised relative to /o/. Since /ɔ/ is identified as /o/ in half of the cases, the best advice is to teach PA learners of American English that /ɔ/ does not exist and that all low back vowels should be pronounced as /ɑ/.

5.4. Discussion

The overall aim of the present study was twofold, and its discussion is partially covered in this section, while the remainder of the discussion is deferred to the following chapter. Meanwhile, the questions and hypotheses of the study can be answered.

No significant differences could be found for PA learners based on the Gender or Context variables in their F1 data, as they showed similar patterns and results, except for the physiological differences that were reflected in shifting the spectral qualities of the female participants to 15% higher formant values than those of the male participants (this shift has been normalized out of the data by within-speaker z-transformation). Additionally, the Gender and Context factors affected F2 values, but only the Vowel effect is of interest here linguistically, and the analysis shows a very strong effect on F2 values. The automatic classification was performed using tokens obtained only in the /hVd/ context, since native control data were only available for this context. The analysis of vowel duration for PA EFL learners shows significant effects of the Vowel factor. Most notably, the AE lax vowels were significantly shorter than the tense vowels, suggesting that PA EFL learners involve vowel duration to differentiate lax vowels from their spectrally adjacent tense competitors. Despite the fact that the tense-lax feature was maintained in PA EFL learners' production, the nonnative vowel durations were shorter overall than the native results for every vowel. Additionally, the Gender and Context factors do not show any significant interaction/effects on vowel duration.

In general, the results of the present chapter show a clear carry-over of Arabic phonetic spectral attributes. The learners' results showed a great deal of overlap in producing AE new vowels. Only vowels with near counterparts (i.e., AE high front/back tense /i/, /æ/, and /u/, respectively) between the two vowel inventories showed a distinct distribution from their spectrally adjacent vowels, as hinted at in section 5.4.1 (but it was also stated that these not confused vowel realizations may or may not concur with the native norm). All other AE vowels were confused in three vowel clusters with partially overlapping centroids and ellipses (/ɪ, e, ɛ/,

/o, ɔ, ɑ, ʌ/, and /ʊ, u/). Mainly, mid-back and mid-front vowels, which are NEW vowels for PA EFL learners, were involved in the confusions. Such overlapping configurations between adjacent new and similar vowels (relative to the PA inventory) are expected to negatively affect the intelligibility of the PA EFL speech production. The PA production results in this relation show a tendency in the learners' responses to keep the AE tense-lax vowels separated, which suggests that EFL learners were able to specify sufficient weight to spectral qualities as they do for temporal ones. Combining the spectral and temporal qualities for PA vowel production yielded a better still configuration of the AE vowels as learners cue duration as a factor related to the tense-lax contrast among AE vowels.

My first question was concerned with the PA learners' production of the AE vowels compared to the AE native speaker results. The informal and short answer is that nonnative vowels differ greatly from the native norm. To answer this question formally, two types of classifications (i.e., LDA and MLRA) were performed twice: once for spectral data (LDA 63.8% and MLRA 65.8% correct identification) and a second time with duration as an added predictor (LDA 74.7% and MLRA 77.4% correct). These results show that the MLRA results for spectral data and when adding duration were slightly better than LDA classifications. The addition of duration as a third predictor increased the percentage of correct vowel identification in the PA results in both types of classification: +11% using LDA and +12% using MLRA.

These initial results only show how well PA EFL learners separate their AE vowel monophthongs. When training the automatic classification algorithms with the native production data results, the nonnative results can be classified by the native model, which allows me to see how the nonnative AE vowels would be classified by native listeners. Here, again, MLRA classification showed an overall better classification for PA results than using LDA. The overall nonnative results in this classification showed a drop in correct vowel identification, but better results were still yielded when adding the (normalized) duration predictor, which makes sense for speakers of L1 with a vowel duration contrast. This confirms my first hypothesis that PA learners have an incorrect (confused) conception of the AE vowel space. My results show that PA EFL speakers are aware of the spectral and temporal features of L2 AE, but they assign more weight to the latter than the former, which partially confirms the second part of my first hypothesis.

My second question asked which AE monophthongs are difficult for PA EFL learners to produce. To answer this question, I first presented the native results and combined the three predictors (F1, F2, Duration). The native speakers produced their vowels in such a way that they were hardly ever confused by the automatic classification algorithms. All vowel types were correctly identified on acoustic grounds at 85% correct or better, with the notable exception, of course, of the *cot-caught* vowel pair. The nonnative results paint a different picture. PA EFL vowel production is confused, especially in the back area of the vowel space with AE NEW vowels. Combining the three predictors for PA results revealed the difficulties in correctly discriminating the vowels yielding confusions such as /u/→/o, ʊ/, /ʊ/→/o, u/, /o/↔/ʊ, ɑ/ (but not /o/↔/u/), /ɑ/↔/ʌ, ɔ, ʊ/ (but not /ɑ/↔/o/, /ɔ/↔/o/, /ʌ/↔/ɑ/, /ɪ/↔/ɛ/, and finally and most importantly, /ɛ/↔/ɪ/, which was confused in 68% of the decisions. It is noted that all these confusions involve NEW AE vowels for PA EFL learners, and confusions with similar vowels between the two languages can be better produced (and automatically identified) in terms of temporal differences. Additionally, there are asymmetrical relationships between confused pairs. The percentage of confusion for members of each confusing pair is not equal and not always reciprocal. For example, /u/ is confused as /o/ 33% of the times but not *vice versa* (/o/→/u/ never reached the 10% confusion criterion to be considered problematic for PA EFL learners), same for /o/→/ɑ/ 14%, but /ɑ/→/o/ never reached the 10% criterion. This suggests an asymmetry such that high vowels are pronounced too low, causing high (close) target vowels to be identified as lower (more open) vowels. Confused pairs were not limited to contrasting members of tense~lax pairs but also occurred within tense vowels (e.g., /u/↔/o/, /o/↔/ɑ/), or within only one lax vowel pair in the front area (/ɪ/↔/ɛ/). Confusions were also noted in the AE native results, in asymmetries favoring high vowels (no misclassifications of high vowels as lower vowels ever reached the 10% criterion needed to be considered problematic). Here, I exclude the more or less symmetrical confusion between the two low-back vowels /ɑ/↔/ɔ/, which is a single merged vowel category in the speech of Californians today. The asymmetrical confusion of /e/→/ɪ/ (15%) is because the diphthongization of /e/ was not measured in the acoustic analysis. A solution would have been to represent the (semi-)diphthongs /e/ and /o/ not by the mean formant values over the entire vowel duration but to extract the formants at 25% and at 75% of the vowel duration and compute the difference between the early and late point in time as a measure of diphthongization (see, for instance, Van Bezooijen & Van Heuven, 2010: 362-363). Interestingly, the /ʌ/→/ʊ/ confusion (10%) coincides with another vowel merger in progress, whereby Californian speakers no longer distinguish between the two (Van Heuven et al., 2020: 124). This confusion did not occur in the vowels of the PA learners; instead /ʌ/ was incorrectly identified as /ɑ/ in the PA results (39%), while the reverse confusion occurred in 24% of the cases. It seems that PA EFL learners spot the differences between new AE back vowels but still confused their constriction place and tend to raise their tongues and open their mouths more when trying to produce NEW back-low vowels in addition to confusing the vowels' duration.

On a second look at the results, tense vowels do not have reciprocal confusions except for the /ɑ/↔/ɔ/ pair, which also occurred in the native results (see above). The PA learners, however, showed smaller confusion percentages for this pair of vowels than the AE natives. Actually, this confusion is the only one in which PA EFL learners showed better performance. Interestingly, PA production of lax vowels is always reciprocal in any confused (acoustically overlapping) pair, except for asymmetrical /ɑ→/ʊ/ (but not /ʊ/→/ɑ/). So, whenever a lax vowel is wrongly identified as a tense vowel, the confusion is reciprocal, but not necessarily so when a tense vowel is confused with a lax vowel. This can be attributed to the fact that duration is the main feature that is phonemically salient for PA learners. These confusions are better identified based on the nonnative duration results in the production task.

The results show a good separation of similar vowels between PA and AE front vowels. Additionally, two pairs of confusions found in AE were not materialized in PA results; those are the /e→ɪ/ and /ʌ→ʊ/ pairs.[38] PA learners also showed better separation for the reciprocal confusion for the pair /ɑ→ɔ/ than the AE natives did. PA EFL results show that the level of difficulty in producing pairwise AE vowel contrasts is: /ɛ→ɪ/, /ɔ→o/, /ʌ→ɑ/, /u→o/, /ʊ→u/, /ɑ→ʌ/, /ɔ→ɑ/, /u→ʊ/, /ʊ→o/, /o→ɑ/, /ɪ→ɛ/, /ɑ→ʊ/, /ɑ→ɔ/ and /o→ʊ/ in a descending order of incorrect identification by LDA trained on native AE vowel tokens (see Table 5.6 for details).

Table 5.6. Incorrect identification (%) of intended vowels produced by PA EFL learners by LDA trained on native AE vowel tokens. The symbol before the arrow denotes the intended vowel, and the symbol following the arrow is the incorrect prediction by the algorithm.

#	Confusion	%	#	Confusion	%
1.	ɛ→ɪ	68	8.	u→ʊ	15
2.	ɔ→o	49	9.	ʊ→o	15
3.	ʌ→ɑ	39	10.	o→ɑ	14
4.	u→o	33	11.	ɪ→ɛ	13
5.	ʊ→u	28	12.	ɑ→ʊ	13
6.	ɑ→ʌ	24	13.	ɑ→ɔ	13
7.	ɔ→ɑ	18	14.	o→ʊ	10

Only /e/, as a NEW AE vowel, did not take part in the confusions of the AE vowel production by PA learners, i.e., did not reach the 10% criterion. These results conform with my second hypothesis and consequently confirm the H2.1 sub-hypothesis concerning the mid vowels, and

[38] By convention, '→' means the left member of the pair was confused (incorrectly identified) as the right member; in pair with '↔' the confusion is reciprocal.

more particularly with the H2.2 sub-hypothesis on how learners may confuse the mid-back vowels. This also concurs with the perception results in Chapter 4.

The overall results on classifying the PA data in accord with AE native results show that similar sounds in the two languages posit less production difficulty for PA EFL learners in comparison with new vowels. Most of the difficulties were statistically found in the NEW vowels, which supports the L2LP model. The exception is the AE /u/ vowel as a near counterpart of PA /u:/ in this case. This finding is supported by SLM and is more likely to be separated by duration. This can be interpreted as learner-individual interlanguages that approximate the native ideal to different degrees.

5.5. Conclusion

This chapter presented the main findings concerning the AE vowels as produced by nonnative PA EFL learners and compared their results to control data obtained from native speakers of AE. The nonnative PA production results show confused alignment of the L2 AE vowel system, especially for the new AE vowels. The results yielded better separation and contrast among the AE front vowels than in the back and mid-back vowels, where most of the confusions were found.

By comparing the nonnative results with the AE native results via two types of automatic classifications, the main confusions in the PA production of AE vowels compared to the native norm surfaced to be in the mid-back area of the AE vowels. Most of the confusions involve new AE vowels, which do not exist in nonnative learners' L1. Only one reciprocal confusion was found within a lax pair, viz. /ɪ/↔/ɛ/, while three confusions were within tense only vowels with one-way direction of confusion /u→o/, /o→ɑ/, and /ɔ→o/, which led me to the conclusion that reciprocal types of confused pairs implicate at least one lax vowel or happen between two lax vowels and, at all times, involve AE new vowels. Despite the somehow sufficient separation of the tense-lax vowels by PA EFL learners, this is the first hint that PA learners L1 interferes with L2 production and blocks their production of new L2 sounds. Moreover, it seems that PA learners are aware of the spectral differences between the AE tense and lax vowels, but more for the former than the latter; whenever a lax vowel is identified as a tense vowel, the relationship is reciprocal, but not necessarily when the tense vowel is confused with a lax vowel. The native results show a similar trend in this regard, with lower confusion percentages and different directionality (nonnative confusions substituted more open vowels, native confusions substituted close(r) vowels). This finding is (almost) certainly caused by L1 interference in L2 production, since Arabic does not have mid and mid-back vowels but still the good separation of tense-lax vowels, especially in the front area, permits the conclusion that PA EFL learners have a sense of AE tenseness and do not depend solely on duration to signal tense-lax contrasts.

My hypotheses concerning the nonnative AE vowel production by PA learners were thus confirmed by the results reported in this chapter. With respect to the first hypothesis, the two automatic classifications show better AE results in vowel production for the natives than for the nonnative. For the second hypothesis, the confusion structures for both groups of participants clearly show these differences, spot the PA exact confusions, and provide error/confusion percentages for each vowel and the directionality of confusions.

CHAPTER SIX

VOWEL PRODUCTION VERSUS PERCEPTUAL REPRESENTATION: EFFECTS OF REGIONAL ARABIC VARIETIES

6.1. Introduction

In this chapter, I try to find answers to two follow-up questions that can be asked after reviewing the results of the perceptual mapping reported in Chapter 4 and the production data presented in Chapter 5. The first question asked was: To what extent does the perceptual representation that an individual EFL learner has of the eleven monophthongs of American English correspond with the way the individual pronounces these vowels? This question can be answered, in principle, by systematically comparing the optimal perceptual targets of the eleven vowels as conceived by the individual EFL learner with the way the same individual produces these target vowels. The data needed to accomplish this comparison are available in the datasets I compiled for chapters 4 and 5.

The second question I want to address is whether the regional variety of the EFL learners' native language, i.e., Arabic, has an effect on the way the learner pronounces the AE monophthongs. I collected data only for Palestinian Arabic learners of English but not for other regional varieties of Arabic. However, numerous varieties of Arabic have been studied in terms of their effects on the pronunciation of English vowels. As a result, I may be able to compare the PA results with similar data (formants F1, F2, and vowel duration) published in the literature.

6.2. Correlation between perception and production

To answer the question of how well the production of the target AE vowels reflects the way the targets are represented in the mind of the EFL learner, I will follow an indirect method. I will first determine, for each EFL learner separately, how closely the vowels in their speech production approximate the native norm. This will allow me to arrange the learners along a continuum from good to poor speakers of EFL. The second step will be to do the same for the learners' perception data. Finally, I can test the hypothesis that the EFL learners' position in terms of production distance correlates well with their position on the perceptual nonnativeness scale.

6.2.1. Simple Euclidean distance

A rather crude but potentially insightful way of ordering the 40 PA participants along a scale of authenticity re. the American native speaker norm would be to compute the Euclidean distance (ED) of each EFL learner from the most typical acoustic values found for the 20 native speakers who served as a reference. As the reference, I take the location of the centroid for all 20 native participants in the hyperspace defined by the 33 acoustic parameters (F1, F2, and duration for 11 vowel types) and then compute the difference for each of the 33 parameters between the native reference point and the location of the individual nonnative participant in the hyperspace. The ED between an individual speaker and the AE reference speaker is then computed as the root-mean-square average of the 33 differences between the individual and the reference centroid:

$$ED = \text{sqrt}\ [\ \sum((D_i - D_{ref})^2) + \sum((F1_i - F1_{ref})^2) + \sum(F2_i - F2_{ref})^2)\ ]\ ,$$

where the index (i) ranges over the 11 vowels /i, ɪ, e, ɛ, æ, ɑ, ʌ, ɔ, o, ʊ, u/ and the term denoted by the subscript (ref) is the reference centroid value.

The measure can be given better diagnostic value if I compute not only the overall ED straightaway but also the distance for each of the 11 target vowels separately. The sum of the 11 separate squared distances is identical to the overall squared mean distance, but by decomposing the distance by vowel, I may learn which vowels contribute more to the overall distance (i.e., degree of nonnativeness) than others. The results of these computations for the production data are shown in Figure 6.1.

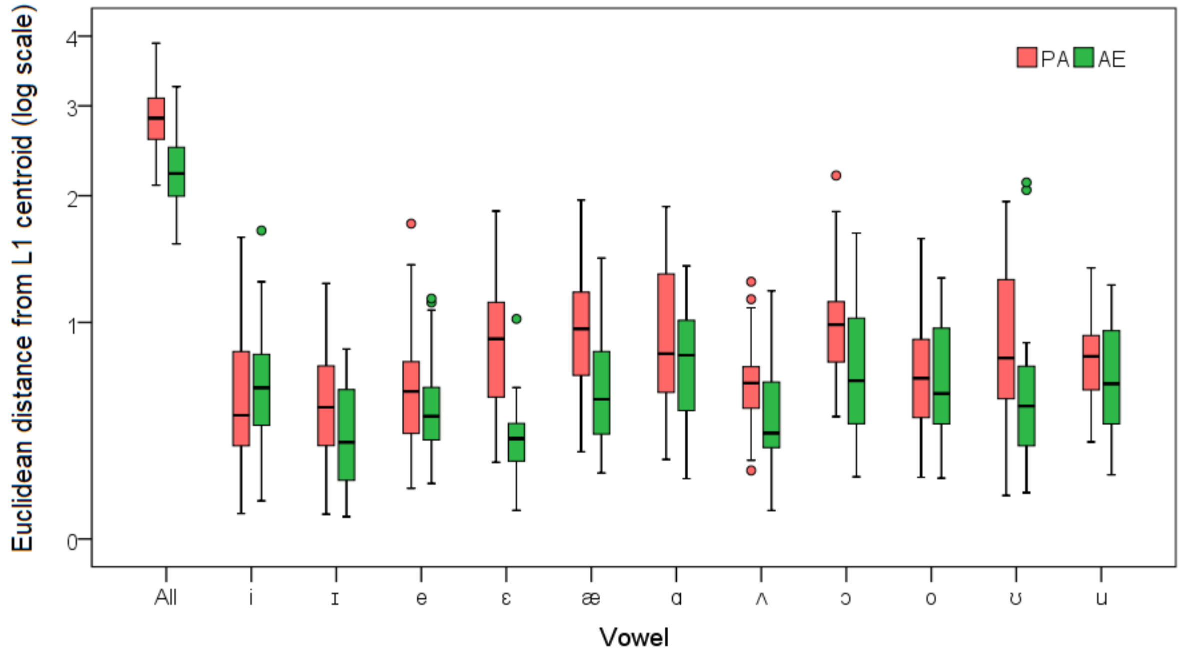

Figure 6.1. Euclidean distance between vowel realizations and American English norm vowel broken down by a group of speakers (40 Palestinian Arabic EFL learners vs 20 American native speakers). Distance is plotted for each of the 11 vowel types separately and for the 11 vowels summed together ('All'). The Euclidean distance scale is logarithmic. Horizontal lines inside the colored boxes are medians. The bottom and top of the colored boxed are the quartiles 1 and 3, respectively. The whiskers extend to the minimum and maximum values found, unless there are outliers (in which case the whisker ends at 1.5 times the length of the rectangle). In the latter case, outliers are explicitly marked.

It is seen in the boxplot in Figure 6.1 that, with one exception (for the vowel /i/), the EFL learners are more remote from the reference centroid than the native speakers for all vowels, as well as for all the vowels added together. Table 6.1 summarizes the statistics for the comparison of the Euclidean distance from the AE norm for the L1 and L2 speaker groups, for each of the vowels separately and for the 11 vowels combined.

Table 6.1. Euclidean distance (after within-speaker z-normalization of acoustic parameters) from norm vowel for L2 (PA) and L1 (AE) speakers of American English, broken down by vowel and for all vowels combined. The difference between L2 and L1 speakers is specified by Δ. The t-value (df = 58) and the p-value for the difference between L2 and L1 are in the bottom two rows.

Vowel	All	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
Eucl. Dist PA	2.858	.584	.562	.641	.869	.957	.933	.666	1.020	.730	.882	.793
Eucl. Dist AE	2.273	.687	.389	.550	.394	.647	.770	.490	.712	.681	.656	.689
Δ	.584	−.103	.172	.091	.475	.310	.162	.177	.309	.049	.226	.105
t(58)	4.891	−1.093	2.230	1.026	5.404	3.227	1.417	2.564	3.169	.553	1.743	1.370
p	**< .001**	.279	**.030**	.309	**< .001**	**.002**	.162	**.013**	**.002**	.582	.087	.176

The overall difference in distance from the norm between L1 and L2 speakers is highly significant for all vowels combined. When broken down by vowel, the difference between L1 and L2 speakers is largest (and with no overlap in interquartile ranges) for the vowel /ɛ/, followed by /æ, ɔ, ʌ, ɪ/ in that order. The other vowels do not differ significantly in terms of their Euclidean distance from the AE norm. This identifies /ɛ/ as the vowel that would be the most difficult to pronounce authentically – as was revealed earlier by the LDA.

The same type of computations can be performed on the perceptual representation data. Figure 6.2 shows the same information as in Figure 6.1 but this time for the perceptual representations.

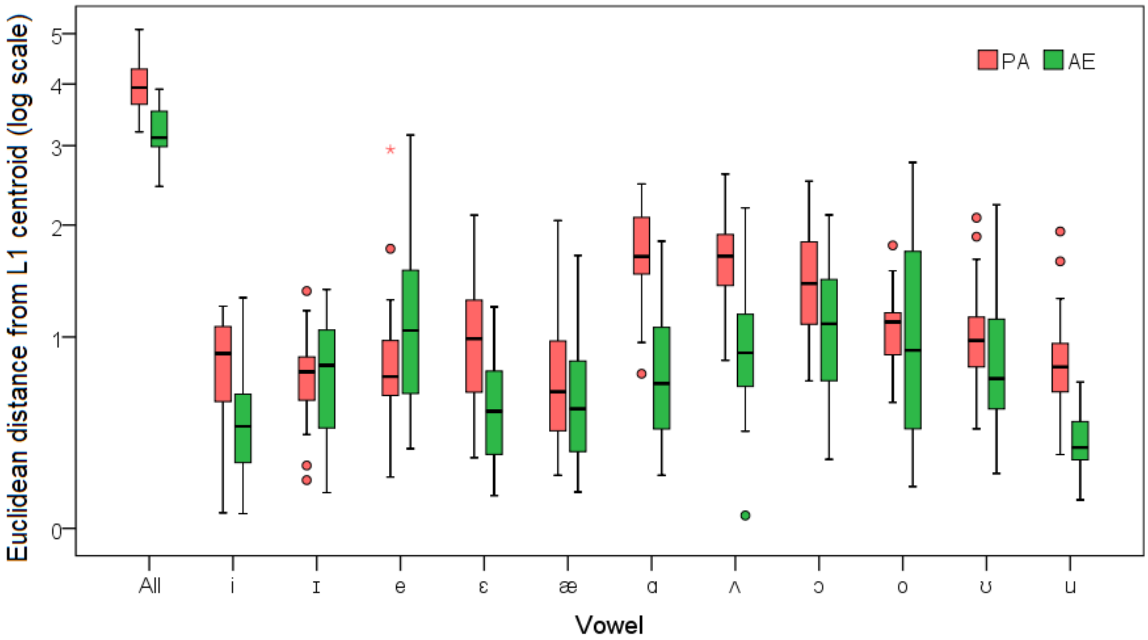

Figure 6.2. Euclidean distance for perceptual representations and AE norms (after within-speaker z-normalization of acoustic parameters) for L2 (PA) and L1 (AE) speakers of American English, broken down by vowel and for all vowels combined. For more information, see Figure 6.3.

Again, it can be seen that the distance of the EFL learners from the native norm is larger than for the native listeners, for the total data and for the majority of the separate vowels. The perceptual representation of the EFL learners deviates most from the native representation for the vowels /ɑ, ʌ/ (with nonoverlapping interquartile ranges), which identifies these as the most compromised pair in regard to perceptual representations. Numerical differences and inferential statistics are presented in Table 6.2. It can be observed in this table that the effect sizes for the perceptual representations tend to be stronger, i.e., with more extreme t-values, than those obtained in the acoustics of the speech production in Table 6.1.

Table 6.2. Euclidean distance for perceptual representations and AE norms (after within-speaker z-normalization of acoustic parameters) for L2 (PA) and L1 (AE) speakers of American English, broken down by vowel and for all vowels combined. For more information, see Table 6.3.

Vowel	All	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
Eucl. Dist PA	3.948	.824	.751	.841	1.002	.736	1.727	1.654	1.461	1.067	1.026	.841
Eucl. Dist AE	3.218	.464	.773	1.197	.580	.632	.772	.968	1.064	1.123	.928	.363
Δ	.731	.359	−.022	−.357	.422	.104	.955	.686	.397	−.056	.098	.478
t(58)	6.630	4.382	−.262	−2.344	3.891	.883	8.107	6.107	3.152	−.402	.811	6.240
p	**< .001**	**< .001**	.795	.023	**< .001**	.381	**< .001**	**< .001**	**.003**	.689	.420	**< .001**

Table 6.2 shows that, overall, the Euclidean distance from the American English norm for the perceptual representation of the vowels is significantly larger for the nonnative PA listeners than for the American native listeners. When broken down by vowel type, the difference is highly significant, and in the predicted direction, for six of the eleven vowel types (indicated in bold face, in yellow cells in the table): /ɑ, u, ʌ, i, ɛ, ɔ/ in descending order of effect size. For /æ, ʊ/, the difference is in the predicted direction but does not reach significance. For the remaining three vowels /ɪ, o, e/ (in red cells), the difference is counter to the expectation but is either insignificant or only just significant (in the case of /e/).

I may conceive of the Euclidean distance from the native norm (centroids obtained for the 20 American native listeners) as a measure of the quality (or authenticity) of the perceptual representations entertained by the EFL learners. Consequently, the 40 PA learners of English can be ordered along a scale of authenticity, both overall and for each of the 11 target vowels separately. This perceptual degree of (non)nativeness can then be correlated with their degree of nativeness obtained from the speech production data for the same individuals. I predict a positive correlation between the perception and the production data. The results of this exercise are shown in Table 6.3.

Table 6.3. Pearson correlation *r* between individual Euclidean distance from American English norm values obtained from speech production and from perceptual representations of 40 Palestinian Arabic learners of English as a foreign language. Probabilities *p* are one-tailed. Values < .05 (marked by *) are significant.

Vowel	r	p	Vowel	r	p
Overall	.118	.234	ɑ	.118	.235
i	−.233	.074	ʌ	.299	.031*
ɪ	−.065	.345	ɔ	−.334	.018*
e	.363	.011*	o	.127	.217
ɛ	−.067	.342	ʊ	−.012	.417
æ	.103	.263	u	.067	.341

As was observed earlier for the comparison of speaker and listener data obtained from the discriminant scores, in Table 6.3, the correlation coefficients are generally low and reveal little systematicity. Three coefficients reach significance, i.e., for the vowels /e/ and /ʌ/ but this is offset by the significant but negative correlation for the vowel /ɔ/.

Figure 6.3 summarizes what was found. It presents the distribution of the overall Euclidean distance of the individuals, i.e., 40 PA learners of English (L2) and 20 American native speakers (L1) from the norm I established for the American English monophthongs.

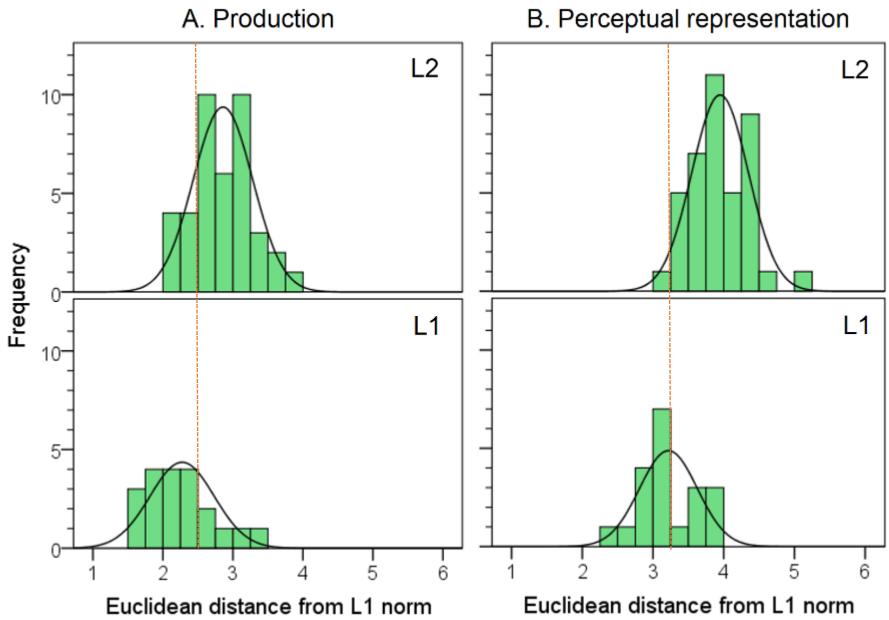

Figure 6.3. Histograms of Euclidean distance from native American norm for 40 Palestinian Arabic learners of English as a foreign language and 20 native speakers of American English. The panels in column A are based on the production data, and those in column B are based on the perceptual representation data. The dotted red line is drawn at the optimal separation between the two groups. The solid line is a normal curve fitted to the distributions obtained.

Figure 6.3 illustrates that the separation between the native and nonnative participants is incomplete, with substantial overlap of the distributions, as is also seen in Figures 6.1-2. Figure 6.3 also shows that the separation by Euclidean distance is better for the perceptual representation than for the active speech production data.

A drawback of the Euclidean distance measure is that it is not sensitive to the direction of differences since it involves a squaring operation that converts every difference to a positive value. I will therefore make a second attempt at establishing the goodness of fit of the AE vowels by an arbitrary speaker (L1 or L2) relative to the AE native norm using a more sophisticated technique, i.e., Linear Discriminant Analysis (LDA). This attempt will be described in the next subsection.

6.2.2. Distance from native norm by LDA

Instead of predicting the type of vowel from its acoustic properties, I may also attempt to predict the native language of the speaker from the acoustic properties of the vowels. Given three properties, i.e., the formant values F1, F2, and the duration, of eleven vowels per speaker, each speaker is characterized by 3 × 11 = 33 parameter values. I use Linear Discriminant Analysis (LDA, Klecka, 1981) as a supervised learning algorithm to classify the set of speakers into two categories, i.e., native speakers of American English vs. PA learners of English as a foreign language. For the production data, I have 20 native speakers of American English and 40 PA learners. Running the LDA algorithm in stepwise mode yields seven parameters (out of the total of 33) that make a significant contribution to the separation of the two groups. The seven parameters together allow a perfect categorization of the 60 speakers. Since there are two categories, only one discriminant function is extracted, which is the weighted sum of the z-normalized values of the subset of the seven parameters. The native speakers receive positive discriminant scores, while the PA learners end up on the negative part of the discriminant scale, with values ranging from −3.53 to .63 for the PA learners and from +1.07 to +5.61 for the native AE speakers. The acoustic variables that contribute significantly to the discriminant function are, in descending order of importance: F1ɛ, F1æ, F1ʌ, F2ɔ, Dɛ, F1ʊ, and F2æ.

I treated the perception data in the same way as explained in the preceding paragraph. The stepwise LDA yielded six acoustic variables that contributed significantly to the optimal split between native and foreign listeners. These were, in descending order of importance: F1ʌ, F1ʊ. Dɑ, De, Du, F2æ. The discriminant scores ranged between −4.26 and +1.41 for the PA nonnatives and between +2.13 and +5.61 for the native listeners. Since there is no overlap between the ranges of the discriminant scores, the classification into native and foreign speakers is perfect. Interestingly, three out of the seven (in the case of production) or six (in the case of perception) significant predictors are the same, albeit in different positions in the rank-order of importance. These are F1ʌ, F1ʊ, and F2æ.

The discriminant score is a composite measure that arranges the members of the two groups under comparison along an optimal distance scale. The closer the discriminant score of a nonnative speaker is to the positive (i.e., native) end of the scale, the more the nonnative speaker behaves like a native speaker. On the grounds of this reasoning, I may consider the discriminant score of the nonnative speaker a measure of the quality ('authenticity') of their pronunciation of the American English vowels or – in the case of the results of the perception test – of their perceptual representation of these vowels.

By way of summary, Figure 6.4 plots the discriminant scores for the perceptual representation and for the acoustic realization of the AE vowels by the native and nonnative participants. Observe how the discriminant scores yield a perfect separation of the native and nonnative participants, clearly better than the Euclidean distance measure plotted in similar fashion in the corresponding Figure 6.3.

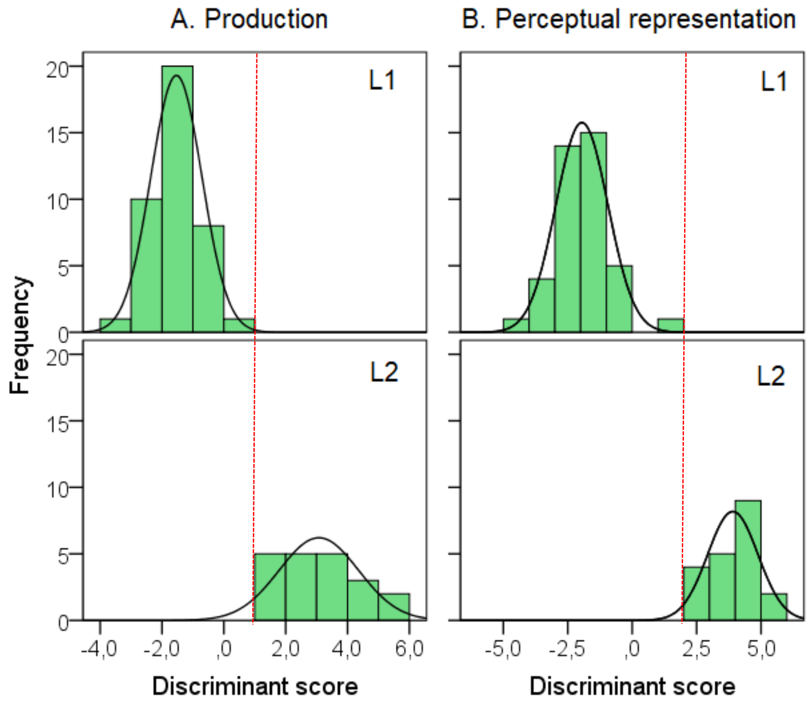

Figure 6.4. Histograms of discriminant scores obtained by LDA for 40 Palestinian Arabic learners of English as a foreign language and 20 native speakers of American English. For more information, see Figure 6.3.

If it is true that a good perceptual representation is a necessary prerequisite for good pronunciation of a nonnative target sound, the discriminant scores obtained for the production and perceptual representation of the 40 EFL learners should be positively correlated. This would mean that speakers with a good (or poor) perceptual representation would also have a good (or poor) active pronunciation of the same vowels, and *vice versa*. Figure 6.5 plots, for each of the PA 20 male and 20 female EFL learners, the discriminant score for the pronunciation of the AE vowels by the EFL learners as a function of the score for their perceptual representation of the same vowels.

Counter to the prediction, however, no strong correlation between the two discriminant scores is found. The correlation is positive, as predicted, but low, and fails to reach significance, r = .140 (p = .194, one-tailed, ins.). Two participants, one male and one female, indicated in the figure by filled markers, seem outliers, in the sense that they obtained a positive discriminant score on one of the two dimensions, indicating atypical closeness to the native speakers in either the vowel production or the perceptual representation. This is unlike the other 38 participants, who have negative discriminant scores for both dimensions. However, even if I omit the two outliers from the data, the correlation remains weak and can be seen as a trend at best, r = .231 (p = .081, one-tailed, ins.).

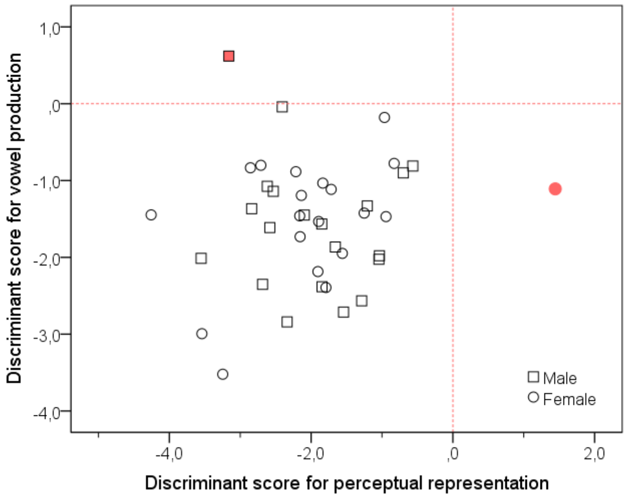

Figure 6.5. Discriminant scores for vowel production as a function of discriminant score for vowel perception, broken down by 20 male vs 20 female PA EFL learners. The correlation is r = .140 (p = .194, one-tailed, ins.) for all 40 data points, and r = .231 (p = .081, one-tailed, ins.) when two atypical participants (with one positive discriminant score, indicated with filled symbols) are omitted.

6.2.3. Interim conclusion

In section 6.2, I tested the hypothesis that the quality of the pronunciation of the English monophthongs produced by an L2 speaker of English should be positively correlated with the quality of the perceptual representation entertained by the same learner. The hypothesis was tested in two different (albeit similar) ways, once on the basis of the individuals' Euclidean distance from the native norms and a second time by establishing and comparing discriminant scores obtained from a Linear Discriminant Analysis classifying speakers and listeners as either native or nonnative on the basis of the formants and duration of the vowels produced and perceived.

In both attempts, a positive correlation was found between production and perception. However, the correlation was too weak to be statistically significant, so that the overall conclusion, based on the present data, must be that the quality (or authenticity) of the individual learner's perceptual representation of the target L2 vowels has little or no predictive power for the same individual's active production of the same vowels (and *vice versa*).

Although no significant correlation could be found between the perceptual representation and the active production of the target vowels at the level of the individual learner, the two domains are obviously correlated at the aggregate level of the native vs nonnative speaker groups. The LDA was able to classify the individual participants into their respective groups (L1 vs L2) without a single error using approximately the same subset of parameters for both the perception and the production data. This shows that the more sophisticated LDA method captures the nonnativeness in the perceptual representation and the acoustic production of the nonnative participants more adequately than merely computing Euclidean distances.

6.3. Effects of regional Arabic variety

6.3.1. Introduction

The AE vowel production results by PA EFL learners show many deviations from the native norms: spectrally, especially in the back area of the vowel triangle, and temporally in the sense that all vowel productions were over-shortened in comparison to the native norms, and last, in maintaining English vowels tense-lax contrast. This section aims to scrutinize several studies of Arabic L1 production of English vowels as reviewed earlier and to compare their findings with mine in light of the three main findings just mentioned (spectrally, temporally, and tense-lax contrast maintenance). Several studies have been reviewed previously in Chapter 5, but their actual results will be discussed in this section of the thesis. I intend to cover not all but a wide spectrum of Arabic varieties and their EFL vowel production to reveal their different effects on the production of English vowels to highlight that every L1 Arabic variety is unique and requires specific recommendations for learning/teaching EFL, especially vowel production.

6.3.2. Spectral differences between different Arabic L1 production of English vowels.

Table 6.4 shows statistics from several L1 Arabic productions of English vowels. The included studies were limited to AE vowels and were produced by L1 Arabic informants in this context and compared to native AE speaker results. Other studies with results on producing British or Australian English vowels were excluded as mentioned earlier. It has been decided so to keep the homogeneity among the compared results as high as possible and to limit the differences to the informant responses of the same AE vowel. Including other L2 varieties production can be done if there is any need to spot differences across L2 English varieties as well but this is not the case in this study. The results of L2 AE vowels in Table 6.4 are compared with the PA results in Table 5.1, which are added at the bottom of Table 6.4 for ease of reference.

Table 6.4. F1, F2 (Hz), and duration (where available, ms) of 11 AE pure vowels as produced by different Arabic L1 varieties, broken down by gender of speaker.

Study	L1 variety	Sex		AE Vowel type										
				/i/	/ɪ/	/e/	/ɛ/	/æ/	/ɑ/	/ɔ/	/o/	/ʊ/	/u/	/ʌ/
Koffi (2021)	Different CAVs	F	F1	351	506	470	527	725	628	538	562	466	420	704
			F2	2595	2104	2315	2036	1958	1303	1169	1242	1349	1182	1508
		M	F1	328	426	470	542	665	685	638	504	440	462	611
			F2	1675	1874	1923	1859	1504	1154	1177	1055	1230	1081	1243
Alqirni (2018)	Saudi	F	F1	425	573	566	524	826	699	760	572	554	465	779
			F2	2656	2066	2277	2490	1930	1520	1520	1223	1231	1163	1488
		M	F1	381	506	510	527	740	699	717	580	571	507	727
			F2	2517	2075	2308	2541	1968	1624	1550	1329	1327	1480	1518
Khalil (2014)	Egyptian	Both	F1	337	460	438	436	625	453	450	441	453	399	672
			F2	2650	2129	2428	2087	2255	1306	1134	1133	1264	1044	1403
Munro (1993)	Different CAVs	Both	F1	319	449	462	608	608	573	N/A	480	431	356	598
			F2	2150	1852	1940	1643	1643	1075	N/A	965	1000	917	1158
			Dur	206	129	225	138	233	159	N/A	230	155	222	146
Farran (2022) This thesis	PA informants	F	F1	411	496	510	507	805	654	607	541	483	451	689
			F2	2664	2230	2436	2170	1796	1496	1403	1384	1598	1411	1579
			Dur	98	57	136	67	114	100	152	113	75	100	68
		M	F1	364	442	453	456	666	530	529	481	411	399	614
			F2	2253	1959	2076	1870	1588	1278	1171	1160	1273	1119	1311
			Dur	93	60	133	68	107	88	133	103	81	95	67

The values in Table 6.4 were psychophysically rescaled to Bark and overlaid by speaker group in an F1 by F2 plot in Figure 6.6A. The vowels produced by the same group of speakers are joined by a convex hull.[39] The convex hull, the vowels joined by it, and the corresponding vowels within the hull are color-coded by speaker group. First, it can be seen that the vowels produced by the five different groups of Arabic EFL learners, although agreeing substantially in the overall shape of the configuration, have different positions in the acoustic space. Typically, the centers of gravity of each vowel system shift along the diagonal from top-right to bottom-left, away from the origin of the plot. This effect should be abstracted from since it is most likely due to linguistically nonrelevant differences in the physical characteristics of the speaker groups, depending on the male-to-female proportion and on the age of the speakers. Female speakers, and younger, physically smaller speakers in general, have smaller resonance cavities and therefore produce higher formant frequencies, which makes their individual vowel configuration shift away from the origin of the F1 by F2 plot. The simplest way to normalize these individual differences out of the data is the z-transformation, when applied to individual speakers, also called the Lobanov (1971) normalization (see also Chapter 5, section 3.4). The effect of the z-transformation is that the center of gravity of the vowel system is uniform at 0

[39] In geometry, the convex hull (or convex envelope or convex closure) of a shape is the smallest convex set that contains it. For a bounded subset of the plane, the convex hull may be visualized as the shape enclosed by a rubber band stretched around the subset.

zF1 and 0 zF2, while deviations for the mean formant frequencies are expressed in standard deviations. Absolute differences in the original centers of gravity and in size of the configurations will disappear, and only the linguistically relevant, relative differences between configurations will remain. Figure 6.6B reproduces the plot in panel A after Lobanov normalization.

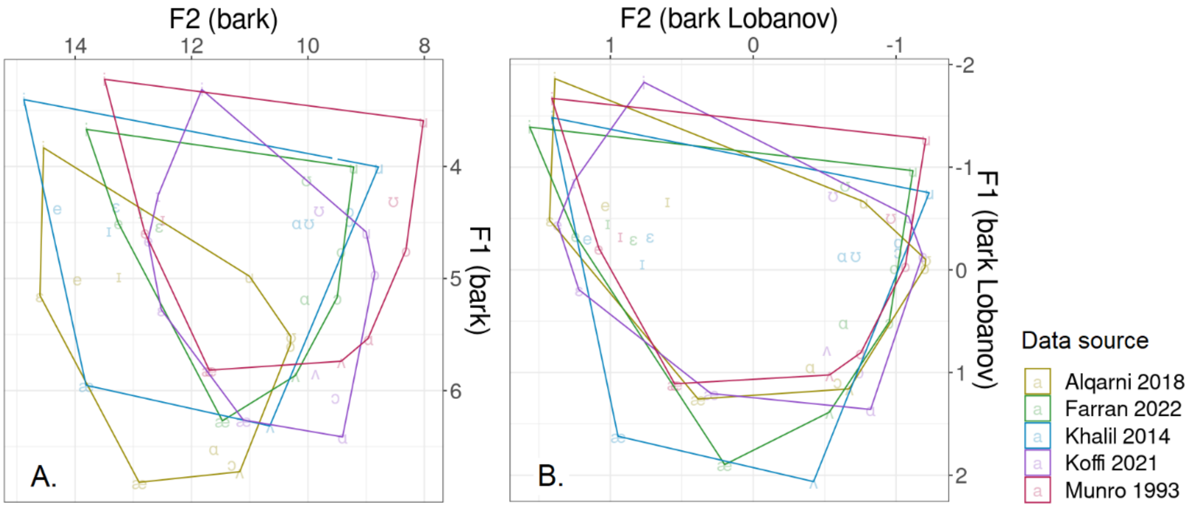

Figure 6.1. A: Mean F1 and F2 values (Bark) of American English pure vowels by Palestinian speakers (black) and by speakers from different Arabic L1 backgrounds (data from Alqarni, 2018; Khalil, 2014; Koffi, 2021; Munro, 1993, for details, see text). The vowels of a specific speaker group are presented as a convex hull, with straight lines joining adjacent vowels along the outer perimeter of the vowel space. Vowels inside a convex hull are indicated by their phonetic symbol only. B: Same as panel A but after Lobanov normalization.

A general observation concerning all L1 Arabic studies is that they all deviate strongly from the native norm, as visualized in Figure 5.7. At the same time, and as could be predicted, they internally share the same pattern concerning sounds with similar counterparts in Arabic. AE /æ/ is the lowest vowel in all F1 values (except Khalil, 2014) and /i/ is the most fronted and highest vowel on the F2 and F1 axes, respectively.[40] My PA results concur with this conclusion as far as the AE /i/ and /æ/ are concerned. These were also the least confused vowels, which may be explained by the existence of similar sounds for them in Arabic yielding positive transfer (as was also reported in these studies). AE /u/ and /ʊ/ were better discriminated by duration. However, even when considering that L1 varieties share very similar vowel inventories (it seems that the back and mid-back vowels are a common issue for all Arabic learners of AE, as discussed in section 5.1.5), the confusions and levels of difficulty in these areas differ from one

[40] An exception to this generalization is Koffi (2021: 80). His vowel /i/ is less front than /ɪ/, but only in the mean data collected from three male speakers from different Arabic backgrounds. Figure 6.6 plots formant values weighted equally for male and female speakers. In the eight female speakers' data /i/ was much more front than /ɪ/. If my means had been weighted proportionally to the number of speakers /i/ would have been more front than /ɪ/.

variety of L1 Arabic to the other. This indicates that each L1 Arabic variety may cause different EFL learning problems based on its own vowel specifics. This process could also further be interpreted as a result of interference by low-level phonetic (articulatory/acoustic) differences between very similar L1s in differently blocking the L2 production in deviating from the L1 vowel inventory sieve. Accordingly, as different patterns of L2 production problems may result depending on the regional variety, tailor-cut recommendations and specific teaching methods are needed to overcome such confusion for each variety.

The confused renditions produced by PA EFL learners of English were found to be significantly different in all three factors (F1, F2, and duration) and more centralized than the native results. These confused productions clarify that EFL learners face problems in determining the target vowels' height, backness, and duration, and it is not sufficient for a vowel to be acoustically separate from adjacent vowels to be correctly produced or intelligible for native (or nonnative) listeners. In the front vowels, for example, the acoustically most confused vowel for PA EFL learners is AE lax /ɛ/ as /ɪ/, the same confusing pair was also reported to be problematic by many L1 Arabic EFL studies but not the same way PA learners did. Khalil (2014) of AE reported the same confusing pair for Egyptian Arabic speakers, but it also involved the /e/ vowel as part of the confusion. Ali (2013) on Sudanese Arabic production of British English, Al-Badawi (2012) on Saudi Arabic speakers of AE reported the same confusions as in Khalil (2014), but not the PA EFL learners. A similar teaching recommendation could be presented for such confusion with some amendments. However, some other problematic English vowels for other L1 varieties were not echoed in my study. Khalil (2014) found the AE /æ/ to be problematic for Egyptian Arabic speakers of AE and its articulation to be higher (closer) than that of /ɛ/, while in this study /æ/ was among the most correctly classified AE vowels (80% correct or more).

Comparing the PA EFL production results and other L1 Arabic studies, Figure 6.6 and Table 6.4 show that PA results are the most centralized among them. This can be linked to interference with PA L1 vowel reality. Studies on the PA L1 vowel inventory show that PA speakers produce their L1 vowels more centrally than speakers of other varieties of Arabic do (Saadeh, 2011).

6.3.3. Temporal differences and preserving tenseness

As mentioned in the literature review, the Arabic language contrasts short and long vowels; therefore, vowel duration is phonemic for L1 Arabic speakers. The educated guess in this situation is that L1 Arabic speakers would employ duration (rather than vowel quality) to realize the AE tense-lax contrast. PA EFL learners in my studies showed good separation between lax and tense vowels in terms of their durations. Nonetheless, their overall durations, both for lax and for tense vowels, were shorter than those of the natives. Additionally, PA vowel durations were shorter than those of all other L1 Arabic varieties. Munro (1993) reported similar findings of shortened AE vowel durations for L1 Arabic speakers compared to the native norm. Unfortunately, other studies do not provide actual duration results in their reports. Figure 6.7 shows the AE vowel duration for PA EFL learners compared to Munro's (1993) and the native AE speakers.

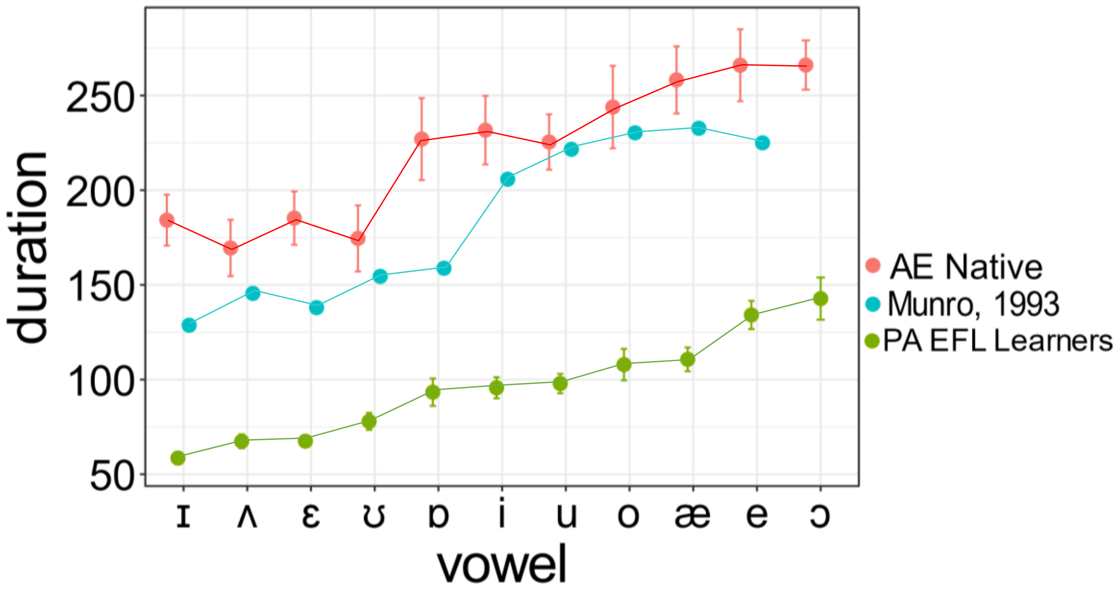

Figure 6.2. AE vowel duration (ms) as produced by PA EFL learners (green), other L1 Arabic speakers (Munro, 1993, blue), and Native AE speakers (Wang & van Heuven, 2006, red). Vowels are ordered by ascending duration of PA EFL learners.

Compared to all other speakers of English, whether native or nonnative, the PA learners produced the shortest duration across the panel for all AE vowels. As briefly mentioned earlier, PA learners were able to separate the durations as shorter for lax vowels and longer for tense vowels, as would the natives do. However, the over-shortening of the vowels might become a source of unintelligibility when hearing PA speakers of AE. They could be simply misheard as saying *pool* and *pull* with the same lax vowel just because both vowels are extremely short, and the difference in duration between them is, actually, very small. Another example would be *bet* and *bit* on the durational differences ground alone.

The alternative, and probably more realistic, hypothesis would be that the native AE listener immediately catches on to the fact that all PA EFL vowels are much shorter than normal and adjusts the duration criteria accordingly. Native listeners are quite flexible in regard to adjusting criteria. They have to do this, for instance, when speakers (habitually) speak fast (or slow) or when they have to interpret vowel qualities produced by male vs female speakers. Adjustment of temporal category boundaries was found to depend on the speaking rate of the immediately following (rather than preceding) sounds (Nooteboom, 1981), while adjusting spectral (quality) boundaries is done on the basis of the quality of immediately preceding vowels (Ladefoged & Broadbent, 1957). According to this view, I would predict that the duration differences in EFL AE are quite effective and will not compromise PA EFL intelligibility. Ultimately, however, future perception experiments will have to give the answer.

The over-shortening of L2 AE vowels by (Palestinian) Arabic EFL speakers can be interpreted and (tentatively) explained in several different ways. In relation to the short-long contrast in Arabic, one issue could be that all Arabic short vowels are symbolized in Arabic orthography by diacritics on the nearest consonant symbol, while long vowels have their own specific base symbols. In English, vowels are always written (diacritics are not part of English orthography), and sometimes a combination of letters or even the same letter can be pronounced differently in different words. PA L2 learners are not taught to mind the duration differences of English vowels, and only by personal efforts would they link durations to the tense-lax contrast in the target vowels. Studies on L1s other than Arabic have reported similar results on EFL vowel shortening (e.g., Munro, 1993).

Another possible explanation might be that the PA EFL learners have an incorrect mental representation of the AE vowel durations. They would then (intuitively) believe that American vowels should be very short and reflect this idea in their active speech productions. This explanation seems highly unlikely given the findings in Chapter 4. The PA EFL learners used the duration cue more effectively than the AE native listeners and had their crossover point from lax/short to tense/long vowels at approximately the same duration threshold as the native listeners (see Figure 4.7).

The third possible explanation would be that the PA learners happened to speak very quickly, thereby reducing the segment durations of all vowels and consonants in their recordings. It is also possible that PA speakers of English apply final devoicing to the stimulus words, which all ended in voiced /d/. This would also affect the preceding vowel, which is near-universally shortened before voiceless stops (e.g., Keating, 1985). A final possibility would be that English is typologically unusual in that it exaggerates the vowel (and sonorant) lengthening before voiced coda obstruents (e.g., Kluender et al., 1988). It seems unlikely that (our) EFL learners are aware of this peculiarity of English. Consequently, they would fail to apply vowel lengthening in their production of the /-Vd/ target words.

These predictions were tested and in a recent follow-up study (Van Heuven & Farran, 2022). It was shown that only the fourth hypothesis was supported by the available data. The PA EFL speakers produced vowel durations before coda /d/ that were equally long as before coda /t/. However, in American English vowels should be at least 100 ms longer before coda /d/ than before coda /t/ (Peterson & Lehiste, 1960: table II). If I add the 100 ms appropriate before /d/ to the PA EFL vowel durations, most of the discrepancy between the Arabic EFL learners and the L1 AE speakers (~125 ms) is eliminated. It is beyond the scope of the present dissertation to present the details of the follow-up study. The interested reader can consult the preliminary report (Van Heuven & Farran, 2022).

CHAPTER SEVEN

GENERAL CONCLUSION, PEDAGOGICAL IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH

7.1. Summary

The present thesis was limited to examining the perception and production of AE monophthongs by PA EFL learners. In Chapter 1, I introduced the background of the Palestinian Arabic (PA) learners of English as a Foreign Language (EFL) to identify their EFL status. Additionally, the pedagogical situation in Palestine was introduced for the readers to reveal the environment where the research studies took place. A bigger picture for the Middle East and North Africa (MENA) region was also given to highlight the fact that the issues approached in this dissertation are problematic for different first language (L1) Arabic speakers in this region of the world. Moreover, the main research questions and broad hypotheses were presented as a starting point for the included research studies that make up the present dissertation and show the organization of the following chapters and how they are connected to each other.

Chapter 2 presented the literature review, which was divided into two sections. The first section consisted of a review of the tackled languages by first explaining the diglossic situation of the learners' L1 Arabic and then by comparing PA quantitative and qualitative features with the phonotactic characteristics/features of American English (AE). This part of Chapter 2 was overlaid with some provisions and hypotheses about the expected issues in acquiring L2 AE segmental and suprasegmental features (vowels, consonants, consonant clusters, and stress) and how the learner's L1 may affect the acquisition process. See Subsections 2.1.4.1–4. The second section of Chapter 2 focused on providing this dissertation with a theoretical framework by reviewing a number of the most updated and well-known L2 learning models, theories, hypotheses, and their most recent versions. The reviewed literature in Chapter 2 shows how different the learners' L1 PA is when compared to L2 AE. The overall importance of this chapter is to guide the dissertation through the experimental studies conducted to try to answer the questions raised in this dissertation in relation to the difficulties L1 PA speakers may face when learning AE as L2 in general and in the perception and production of AE vowels in particular. The main methodology of this dissertation is based on assimilation, identification (perception), and production testing in three consecutive studies in which informants were nonnative PA EFL learners and native AE listeners/speakers. The three studies of this dissertation have been presented in Chapters 3-5.

Chapter 3 addressed PA EFL learners' patterns of assimilation of AE L2 vowels to PA vowels via a test method developed within the framework of the Perceptual Assimilation Model (PAM). The results of this chapter showed massive confusion of the AE monophthongs by PA EFL learners, predicting identification errors that PA learners might be vulnerable to in perceiving AE vowels in Chapter 4. In contrast, the fulfillment of the PAM predictions in the identification experiment in Chapter 4 can be seen as an indication that PAM manages to predict EFL learning problems.

Chapter 4 incorporated a perception experiment with synthesized vowels, which had to be identified as tokens of the AE monophthongs by native AE and PA EFL listeners with the aim of revealing differences in the mental representation of the vowel space between native AE listeners and nonnative PA learners of AE and how the two differ by mapping out and comparing the nonnative identification of the synthesized vowels with the modal/optimal mapping by the natives at the individual level. The results of Chapter 4 show, first, that PA EFL learners differentiate between AE tense and lax vowels but mainly rely on duration differences to discriminate between spectrally adjacent members of a tense-lax vowel pair (see section 4.3.3). This result can be interpreted as a negative transfer from the PA vowel inventory with its 3 vowel pairs that are mainly considered to contrast in duration only, while the quality of each short-long vowel pair is thought to be the exact same. Additionally, the results of this chapter showed that only near equivalent vowels between L1 and L2 were correctly identified by PA EFL learners. Regarding the quality differences, PA EFL learners showed huge deviations from and substantial confusion of the AE native norms and the modal responses, as seen in section 4.3.4 and Figure 4.8 in particular. Additionally, the most confused deviations of PA learners in perceiving AE vowels were represented in Figure 4.9. Table 4.4 verifies my results and illustrates the specified weight attached to discriminating AE vowels in relation to duration vs. quality. In short, the results concur with the predictions derived from the PAM results in thus far as the AE high tense vowels are concerned. The agreement of PA perception choices with the native AE results, in general, did not reach 34%, except for the high tense vowels, where it reached only 50%. The PA EFL learners' perception of AE vowel duration is highly congruent with the native norms. Meanwhile, the correlation for perceiving the vowel quality between the natives and the nonnatives is weak. Here, the results are also in agreement with the PAM predictions concerning the hypotheses on the importance of duration as a primary source to differentiate between nonnative AE tense-lax vowels and the how the mid-section vowels are confused because they do not exist in the learners' L1.

In Chapter 5, EFL learners produced the AE pure vowels, which were then compared with control data obtained from native AE speakers, with the aim of determining how the production of AE monophthongs by nonnative PA learners differs from the native norm in light of the perception results in Chapter 4 to reveal whether there is a correlation between nonnative perception and the production of AE vowels. Among the chapter's findings were the insufficient separation (overlapping dispersion ellipses) between adjacent vowels in the IPA vowel chart and the strongly overlapping production of AE vowels that have no counterpart in the learners' L1, mainly in the central part of the vowel space. Nonnative learners did not express awareness of spectral properties and did not associate them with their temporal properties. Among the findings, the good separation in perceiving the tense-lax vowels in Chapter 4 is echoed in the production results by the PA EFL learners. Interestingly, and unexplained thus far, all the AE vowels are over-shortened in the speech production by the PA EFL learners, while relative duration differences largely mimic those of native AE speakers. Moreover, no tendency to shorten AE vowels was found in the corresponding perception data.

Chapter 6 synthesized the three research studies and compared the learners' perception and production in an integrated fashion to determine whether a good perceptual representation is a necessary prerequisite for a good pronunciation of a nonnative target sound. Counter to my predictions, the results did not report a strong correlation between the learners' perception and production of AE vowels at the individual level. Next, the chapter compared the production data of the PA informants with speakers of other regional varieties of Arabic learners of AE in terms of three phonological features: vowel formants, duration, and the generalized phenomenon of over-shortening L2 vowel durations compared to L2 native norms.

The remainder of this chapter recapitulates the main objectives/questions underpinning the research in the present dissertation in Section 7.2. The main research questions will be recalled one by one in separate sections to highlight the supporting results that were provided for them and present applicable solutions and pedagogical implications thereof, while the accumulative answers for each study try to provide a better understanding of the overall research question, which was how the properties of the PA L1 sound system may influence the acquisition of AE as an L2 at the segmental level. The study limitations are presented in Section 7.3. Next, Section 7.4 discusses some pedagogical implications for the EFL curriculum developers, language educators, and learners. Finally, section 7.5 summarizes my future plans and ongoing studies.

7.2. Recapitulating the main objectives of the dissertation

With regard to the study questions and as part of my general hypotheses for study 1, I anticipated that L2 new/additional features not available in learners' L1 would reveal potential assimilation problems, such as the mid vowels in the AE inventory, and then specific movements or gravitations were hypothesized for each AE vowel toward some PA vowel based on the shared or common attributes for each vowel in each inventory. Accordingly, the overall question for study 1 was: How is the PA vowel system predicted to affect the perception of AE pure vowels? A number of specific questions resulted from the main question.

Chapter 3 (Study 1). The first group of questions specified for the first study attempted to reveal how the PA listeners assimilate the AE vowels to their L1 PA by employing PAM and its predictions on discrimination and identification patterns. The PA vowel inventory has fewer vowels than the AE vowel inventory with very few near counterparts, as discussed in the literature review chapter, which would naturally lead to a mismatch between the two inventories and thus cause L2 perception problems at the vowel level. Therefore, the questions for the L1-L2 perception study (Study 1) were meant to reveal how listeners would (potentially) warp the perception of AE pure vowels to fit the PA vowel system and predict perception problems for L2 AE vowels. After reviewing relevant studies on L1 Arabic perception of English as L2, the study questions for the PA L1 situation were as follows:

1- Which AE vowels constitute a (potential) perception problem for PA EFL learners?

The results of the first study provided general assimilation scenarios (based on PAM) to the L1 vowels they assimilate to, which allows the prediction of confused perception of the AE vowels in pairs or triplets. These results were based on a statistical analysis of the informants' responses that counted for their choices of L1 and goodness of fit. The results yielded 55 possible contrasts that PA EFL learners can make of AE L2 vowels. The results show that PA EFL learners mapped eight vowel pairs onto the same single vowels in PA, seven of which were in an SC scenario, i.e., the members of each pair were considered equally good exemplars of a single PA vowel: (1) *heed* /i/ ~ *hayed* /e/, (2) *hid* /ɪ/ ~ *head* /ɛ/, (3) *hawed* /ɔ/ ~ *hod* /ɑ/, (4) *hawed* /ɔ/ ~ *hoed* /o/, (5) *hawed* /ɔ/ ~ *who'd* /u/, (6) *hood* /ʊ/ ~ *hud* /ʌ/, and (7) *hoed* /o/ ~ *who'd* /u/). One further pair yielded a CG scenario, in which both AE vowels map onto a single PA vowel but with a difference in goodness between them: (8) *had* /æ/ ~ *hod* /ɑ/) as detailed further in Section 3.3, Tables 3.1-3, and Figure 3.2. These were then predicted to be the most confused AE vowel pairs for PA EFL learners; the rest (47 contrast) are hypothesized to not constitute a learning problem. This part answers the first question of this study.

The second main question for this study consisted of two related parts:

2- (a) How perceptually sensitive are PA learners to AE vowel duration and/or quality? And (b) Which AE vowels are predicted to be the most difficult to perceive?

With respect to question 2(a), I found that the mid-high tense AE vowels /e, i/ project onto PA /i:/ and /o, u/ onto PA /u:/, and will therefore be difficult to differentiate from the AE high tense vowels (PAM pairs 1 and 7 above). The lax vowels /ɪ, ʊ/ may assimilate to PA short /i/ and /u/, respectively, which may also be the case for AE /ɛ/ and /ʌ/ (PAM pairs 2 and 6). These sub-hypotheses, specific for each AE vowel, were fully confirmed by the yielded results in Section 4.3.

As predicted, the rest of the vowels (the mid and back vowels) created massive confusion, where /o, ɔ, ɑ/ projected onto long PA /u:/ (PAM pairs 3, 4, and 5). Additionally, an exception was hypothesized for the /ʌ/ and /ɛ/ to project onto short PA /a/. In that case, AE /ʌ/ and /ʊ/ would form a TC contrast (see Figure 2.3 in Subsection 2.1.4.1). My results show that PA short /a/ did not qualify as an assimilation category for any of the 11 AE vowels, i.e., no English vowel comes close enough to it. That ruled out the last part concerning the exception of /ʌ/ and /ɛ/ projection onto short PA /a/ that would lead to well-separated assimilation between /ʌ/ and /ʊ/ as members of a TC contrast. On the contrary, /ʌ/ and /ʊ/ (PAM pair 6) were found to fall in an SC scenario as both were assimilated to PA /u/, and lasting learning problems are expected for this type of scenario, confirming the more general hypothesis in this part. Concerning the first part of the hypothesis about the (mid)low vowels, the results in Section 3.3, especially Tables 3.1-3, show that only AE /æ/ and /ɑ/ (PAM pair 8) assimilated to PA /a:/ in a CG scenario, which is part of what I predicted for the AE low vowels. However, PA EFL learners also found /ɔ/ and /ɑ/ to assimilate to PA /u:/, which falsifies this part of the original hypothesis.

The last hypothesis in the first study concerned the perception of vowel duration. I hypothesized that the PA learners would most likely depend on the temporal cue more than the quality of a vowel to discriminate between the lax vs tense members of spectrally adjacent vowel pairs. Vowel duration is predicted to be not problematic and will be the least confused feature by PA learners of English, in the sense that members of the Same Category AE vowel pairs were either both tense or both lax; hybrid tense – lax AE vowel pairs are never assimilated to the same single vowel type in PA. Instead, confused tense vowel pairs assimilated to PA long vowels only and confused lax vowel pairs assimilated to PA short vowels only, as seen by eyeballing Figure 3.2, which is supported by the results in Table 3.4. These findings answer to question 2a. It is also clear from these results that PA learners have a sense of the AE tense–lax contrast but relate it primarily to vowel duration.

Another multifaceted finding in this relation is that the potentially problematic AE vowel pairs or triplets contain spectrally AE adjacent vowels only and involve one or both vowels that are not available in the learners' L1, which happen to assimilate to PA vowels in spectrally the same vowel area in the vowel space. This finding reveals several things. First, it explains the learners' divergent goodness ratings in the PAM results. Second, it hints at PA EFL learners' assimilation process of L2 vowels based on the vowel quality in addition to vowel duration, which was basically intact. Furthermore, it reflects on the notion that learners with a small L1 vowel inventory do not necessarily involve unused areas in the vowel space in their perception. The results show that the six vowels of PA divide up the vowel space exhaustively (i.e., there is no empty space). This interpretation is supported by the learner's high goodness rating for the L2 vowels that are unavailable in their L1 (rated as good or fair exemplars of L1 vowels (see the goodness ratings in Table 3.2).

As for question 2b, I opted to provide a specific ordering for the level of difficulty for the confused vowels. Therefore, the decision was made to employ the fit index as a measure that involves both the absolute numbers of the informants' responses for each vowel and the yielded goodness of fit for each vowel to rank order the confused vowels based on their difficulty weighted by the just-mentioned values for each vowel. The fit index spotted the confused AE vowels, rank-ordered them with a comparable index to another/other member(s) of the confused pairs or triplets, and showed to which PA vowel each AE vowel assimilated. Accordingly, the fit index results in Table 3.5 in Section 3.3 answered question 2b in this study.

The third question was as follows:

3- How do the results of this research align with the results of similar studies on other CAVs learners of EFL?

The studies reviewed in Chapter 3 on other varieties of L1 Arabic provided results that were in line with the results for the PA EFL learners sometimes and diverted other times. Specific discussion for each L2 vowel is provided at the end of Section 3.3 and the discussion section in Chapter 3 (Section 3.4). A more detailed discussion is provided in the comparison with Yavaş's (2011) results in section 3.4.1 by comparing the results in Figure 3.2 with Yavaş's results in Figure 3.3.

Chapter 4 (Study 2). The purpose of study 2 in Chapter 4 was to test the PAM predictions and assimilation categories from the previous study for the PA EFL learners on the perceptual representation of the full set of AE monophthongs using synthesized vowels that cover the human vowel triangle to be identified as phonemic AE vowels. By doing so, I was able to define the target positions of the AE vowels as entertained by PA EFL learners, to evaluate the relative importance of spectral vs temporal properties of the vowel categories, and to estimate the fuzziness of the vowel categories. This in turn allowed me to identify which AE contrasts are difficult to discriminate by PA EFL learners and to compare their results with native AE results for the same task as well as to reveal the perceptual mappings and the effects of spectral and temporal features of a vowel as a source of discrimination difficulty. The general objective of this study was to reveal the PA learners' perceptual sensitivity toward AE vowel features spectrally and temporally as well as to determine which are the most difficult AE vowels for PA EFL learners to identify. Therefore, a set of 3 sub-questions were formulated as follows:

The first two questions in chapter 4 asked were

(i) How do Palestinian Arabic listeners conceive of the American English vowel space? and
(ii) How do American native listeners conceive of their own vowel space?

The analysis performed on the acquired data showed that nonnative learners struggled with their mental conception of the AE vowel space in general. In total, their modal responses were not high; the modal response average for the set of stimuli reached only 51.8%, while two AE vowels (/ɛ/ and /ɔ/) were never chosen at all as modal response categories. Their performance did not measure up against the modal responses of the native participants and indicated uncertainty on the part of the EFL learners regarding how to identify the AE vowels. Their confusion network in Figure 4.8B revealed more confused connections compared to the native results for the predicted to be problematic pairs of AE vowels, especially in the mid and back sections of the vowel space. As predicted from the results in the previous chapter, a similar trend was found for the confused vowels, especially for those that do not have a counterpart in learners' L1. Serious mapping misconception resulted for the members of the confused pairs of vowels in the PAM test but with different AE vowels than those in the original PAM pairs in the Same Category or Category Goodness scenarios. These findings called attention to the importance of the temporal attribute in correctly organizing the perceptual mapping in the mind of the nonnative learners in comparison to the native norm. Also, with regard to L1 interference as a key factor in warping the EFL learners' perceptual mapping. The native results showed a better and more clearly defined perception of the set of stimuli with 59.7% modal response average, but they did not perceive (/e, o/), which was attributed to the lack of diphthongization of the synthesized vowel set. See Section 4.3 and, more specifically, Subsection 4.3.4 for further details for both groups of participants.

The third question asks:

(iii) How does the mental representation of vowel sounds differ between AE native listeners and PA learners in terms of vowel quality and duration, and how do quality and duration interact or trade?

To answer this question, I quantified the differences between the nonnative PA results and the native AE vowels in order to reveal how the PA perception of AE vowels diverges from the native AE norm. In Section 4.3.4, Figure 4.9 illustrated the differences between the nonnative and native results. This final step showed that the agreement of the PA decisions with the modal AE decisions is only 33.8%. At the level of the single response category, agreement with the AE norm exceeded 50% only for two response categories, namely, /i/ (51.4%) and /u/ (51.6%). Agreement with the native modal response was much lower for the other nine vowel types and never reached 35%. This is strong evidence that the PA EFL learner's perceptual mapping of the AE vowel space is significantly different from that of the AE natives.

Chapter 5 (Study 3). In Chapter 5, the same group of nonnative PA speakers as before produced the 11 AE vowels in carrier words and phrases. Their recordings were compared with those produced by native AE control speakers. It was anticipated that learners' incorrect perceptual representation of the AE vowels would be reflected in their production. Therefore, the overall purpose of Chapter 5 was to determine how the PA production of AE vowels differs from the native norm and to see whether the deviations are similar to what was observed in the perceptual representation. Therefore, two questions and several hypotheses were raised, which were provisionally answered in Sections 5.3-4 and are summarized below. Two additional questions were prompted by the results, and their discussion was deferred to Chapter 6.

The first question asked:

(i) How do PA speakers produce the AE vowels in terms of quality and quantity in comparison to the native AE participants' production?

Different types of informal and formal statistical analyses were performed on the data. First, the PA learners did not show statistical differences based on gender in the spectral constellation of the AE vowels, so the results were aggregated over boys and girls. Additionally, for temporal properties, PA EFL learners show great consistency for the vowel sequence based on duration and good temporal separation between the tense and lax AE vowels. These results were attributed to the fact that the nonnative learners' L1 employs duration as a phonemic feature. Nonetheless, it was mentioned that nonnative speakers with no length contrast in their L1 were reported to behave similarly to the PA learners of AE with respect to the AE tense-lax contrast (e.g., Perwitasari, 2019).

Formally, I tested the effects of vowel, gender, and context factors on the learners' F1, F2, and duration. The results of these statistical tests revealed how each vowel attributed is affected by each factor. See Section 5.3.2 for more details.

The same informal and formal analyses were performed for the native AE results in Section 5.3.3. Preliminary differences were noted for the nonnative results as they depart from the native norms both in spectral and temporal properties. Accordingly, a multivariate analysis was performed to pursue these findings further in Section 5.3.4 using two different classification algorithms, i.e., Linear Discriminant Analysis (LDA, Klecka, 1981) and Multinomial Logistic Regression Analysis (MLRA, Hosmer & Lemeshow, 1989), once for the spectral parameters only and the second time with vowel duration added to isolate the weight of vowel duration (see Table 5.4). Moreover, Table 5.4 shows that adding vowel duration as a third predictor of vowel category in the automatic classification of the 11 AE monophthongs improves the correct classification by 11 percentage points (from 65 to 76% correct) when the classifiers are trained and tested on the nonnative tokens. When the classifiers are trained on the native tokens and then tested on the PA EFL tokens (using the classifiers as substitutes for human AE listeners), adding vowel duration has a smaller effect (Table 5.5). The native speakers gain 7 or 9 points (depending on the classifier), while the nonnatives gain 9 points in the case of LDA but only 4 points in the MLRA. These results indicate that vowel duration is used by the PA learners of English in a systematic way to differentiate between vowels in AE but in a way that will not be understood by native AE listeners. The PA duration differences will quite probably help PA speakers of AE to identify the PA-accented AE vowels as intended by the nonnative speakers. This would be an example of the shared interlanguage speech intelligibility benefit (Shared ISIB, Bent and Bradlow, 2003; Wang and Van Heuven, 2015). More details are provided in Subsection 5.3.4.2.

The second question for this study asked:

(ii) Which AE monophthongs are difficult to produce by PA EFL learners?

The result of the automatic classification showed a clear confusion of the AE vowels produced by the nonnative learners. Nonetheless, the confusions were specific for specified clusters of adjacent vowels only with no front-back vowel confusion. Figure 5.10 shows the nonnative vowel confusion, which can be compared to the native results in the same figure. It is shown that the most significant confusion happened between two front lax vowels (/ɪ, ɛ/).

Chapter 6. Based on the results of the identification and production tasks for the nonnative learners, I tried to answer two follow-up questions in Chapter 6. First, I asked to what extent the perceptual representation that an individual EFL learner has of the eleven monophthongs of American English corresponds with the way the individual pronounces these vowels.

Chapter 6 shows that there is no clear correlation between perception and production at the level of the individual learner. I established a rank order among the 40 participants separately for the perception task and for the production task. The greater the participant performed the task like a native listener of speaker, the better the individual's ranking. Then, I correlated the perceptual ranking and the production ranking and found only a weak correlation (just not significant, merely a trend with a p-value between .05 and .10).

However, a better correlation can be established when I compare the two datasets and count confusions. As discussed earlier in Chapter 4, I considered every confusion that occurred in more than 20% of the responses to a particular vowel type (where confusion was defined as "deviating from the modal native response vowel") as serious. These are the reddish cells in Table 4.3. Less frequent confusions were not highlighted in this table. The green cells are the correct responses. These should not be included because their contents are predictable from the error percentages. I reproduce Table 4.3 below as Table 7.1, listing the percentages only while omitting the marginals.

Table 7.1. Percentage of vowels identified by PA EFL learners in 86 synthesized vowel sounds in /mVf/ context. The correct response is defined as the modal response category as identified by 20 native AE listeners. Correct classifications are in bold face in green cells along the main diagonal. Error percentages ≥ 10 are in red cells.

		All observed nonnative responses										
		i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
Modal native response vowel	i	**51.4**	31.2	2.1	11.3	.3	.6	.9		.9	.9	.3
	ɪ	3.3	**34.1**	2.4	9.8	1.6	7.3	2.0	1.2	7.7	13.8	16.7
	e											
	ɛ	4.0	27.4	32.6	**17.1**	2.7	2.1	2.4	4.3	2.7	3.7	.9
	æ	.2	5.2	15.8	3.4	**27.3**	1.5	36.3	5.0	1.8	2.9	.7
	ɑ		.6	4.3		2.4	**19.5**	5.5	9.8	33.5	14.0	10.4
	ʌ	.3	1.4	2.4	1.4	8.4	21.6	**22.3**	8.7	12.2	16.4	4.9
	ɔ			6.7	.6	11.6	12.8	49.4	**5.5**	3.7	9.8	
	o		.6		1.2	1.2	22.0	.6	7.9	**33.5**	19.5	13.4
	ʊ	.8	4.1	6.1	4.1		24.4	2.0	6.9	21.5	**19.5**	10.6
	u	1.1	2.2	.9	1.1	.3	10.4	1.7	1.2	11.3	18.1	**51.6**

Table 7.2 presents the information found for the production data by running the LDA trained on native AE speakers to classify the vowels as produced by the EFL learners. The complete table is not in chapter 5 but can be found in Appendix 5.5.

Table 7.2. Percentage of vowels produced by PA EFL learners, as identified by an LDA trained on native AE vowel tokens (20 tokens per type). Error percentages ≥ 5 are in red cells. See Table 7.1 for more information.

Modal native response vowel	All observed nonnative responses										
	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
i	**82.2**	11.0	6.8								
ɪ	1.8	**84.2**	2.6	11.4							
e	.9	5.5	**89.0**	1.8	2.8						
ɛ	4.1	65.2	2.3	**28.1**	.5						
æ				1.7	**93.2**	4.2		.8			
ɑ					.9	**40.0**	13.9	16.5	9.6	15.7	3.5
ʌ				.9	.9	45.9	**46.8**	2.7	1.8	.9	
ɔ					1.0	12.7	1.0	**45.1**	40.2		
o			1.0			21.4	6.8	1.0	**57.3**	7.8	4.9
ʊ		2.9						1.0	19.2	**49.0**	27.9
u						.9			42.1	8.8	**48.2**

Generally, the percentage of correctly identified vowels is higher in the production data. Therefore, the error/confusion percentages are smaller. Accordingly, I considered confusions that occurred in 5% or more as serious. Even then, the number of red cells in the production data is smaller than that in the perception data.

In all, there are (11 × 10) = 110 cells in each table (11 squared minus the correct cells along the main diagonal). I now ask to what extent the two tables are alike in the location of the red cells. If there is congruence, I should find that a red cell in the production table matches a red cell in the perception table. There are 20 red cells in the production table (versus 90 white cells). By the same token, there are 29 red cells in the perception table (against 81 white cells). Contingency is computed as can be seen below in Table 7.3.

Table 7.3. Contingency table of the number of serious confusions in the production of AE vowels by PA EFL learners against serious confusion in their perceptual representation.

Confused in production	Confused in perceptual representation		Total
	No	Yes	
No	**76**	14	90
Yes	5	**15**	20
Total	81	29	110

Out of the 20 strong confusions I find in the production data, 15 (75%) are also found in the perception data. However, out of the 90 non-confused vowel pairs in the production data, only 14 (16%) are strong confusions in the perception results. This is at least a moderate association between the confusion structure in the perception and production data when considered by vowel type, $\varphi = .520$, $\chi^2(2) = 29.8$ ($p < .001$).

The second question in Chapter 6 inspects whether the regional variety of the EFL learners' native language, i.e., Arabic, has an effect on the way the learner pronounces the AE monophthongs. The analysis was based on the EFL production data retrieved from studies on regional varieties of Arabic other than the PA variety. Moreover, by providing findings in this chapter for the three different studies in Chapters 3-5 while focusing on the learners' perception and production relationship, I was able to explicitly discuss how the PA EFL results compare to other varieties of L1 Arabic, which can strengthen my conclusion on how learners' different L1s can affect their acquisition of AE differently even if all these varieties are seen as representatives of an umbrella language, i.e., Arabic.

The way the PA EFL learners perceive and pronounce the vowels of AE largely converges with what is reported in the literature on L2 acquisition of (American) English vowels by learners from Arabic backgrounds other than Palestine. However, there are some discrepancies observed in the PA learners from the overall picture: first and foremost, the over-shorting of vowel duration compared to other L1 regional varieties of Arabic reported in Munro (1993). Despite the fact that PA EFL learners preserved a good contrast between the AE tense and lax vowels, the overall durations were shorter than any vowel duration reported in Munro (1993) (and shorter than the vowel durations found for the native speakers). Spectrally, PA EFL results were found to be more centralized (shrinking the vowel space) than those reported for the other Arabic L2 studies. While the results show congruence in the location of the AE vowels for some Arabic varieties when compared to the PA EFL results in this dissertation, considerable shifts were found for other vowels, and no other L1 Arabic variety shows the same spectral patterns as the ones reported for the PA EFL learners in this dissertation. This can be attributed to the fact that L1 regional Arabic varieties differ from each other and accordingly affect learners' L2 production differently. This, in turn, leads to the conclusion that the teaching of the correct pronunciation of AE vowels should differ and may use different drills and exercises, depending on the regional Arabic variety of the EFL learner, to accommodate the different native-language interference effects.

7.3. Limitations

There are some limitations at different levels in the dissertation. First, this dissertation, ideally, should be a longitudinal study, with repeated testing over a number of years. The limitation in this case is that the PA informants are young adolescents and do not have much time to act as informants in a longitudinal study. The importance of consecutive testing at different points in time lies in its ability to detect the development in learners' perception and production over time so that I may break into the chain of cause and effect. My current findings indicate that a good perceptual representation of the L2 vowels correlates with good production at the aggregate level (but not at the level of the individual learner), i.e., L2 vowels that were confused a lot by PA EFL learners in Chapter 4 are also those that were confused a lot in Chapter 5 by the same participants. However, this does not prove that good perception is a prerequisite for good production at the individual level. For the latter causal relationship, I should be able to show that good perception precedes good production, which can only be done if I measure the students' performance at multiple, different points in time.

I wanted to dig deep into the L1 PA varieties and test the potentially different effects of rural vs. urban dialects on AE perception and production, as each dialect has exclusive features that can affect the acquisition of EFL differently, especially at the consonant level (see the literature review in Chapter 2). Due to time constraints on this PhD project and data collection during the COVID-19 pandemic, I was not able to collect a representative sample from speakers of each dialect; therefore, these plans are shelved for future research. In this relation, I also wanted to test the tasks with a control group of absolute novice PA listeners with zero experience with English to check whether the PAM results would be any different from those of the available participants (who had several years of English lessons).

Although I did argue that pure vowels are a primary source of a foreign accent (compared with diphthongs in case L1 and L2 have diphthongs), the pronunciation of consonants and, more particularly, consonant clusters may well be a more important source of unintelligibility. In this PhD project, the recorded data and materials are far more than what are actually analysed, and the analyzed materials are more than what is actually used in this dissertation. Among the unused data are recordings of continuous speech (The North Wind and The Sun fable, both in English and Arabic), English consonant clusters, and single consonants in onset and coda positions. Including the analysis of these materials would have helped more in revealing the bigger picture on how L1 PA features affect the acquisition of segmental and suprasegmental L2 AE features for PA learners.[41]

7.4. Pedagogical implications

The studies conducted in the present dissertation have shown that the acquisition of English as a foreign language by PA learners is compromised both in the perception and production of AE

[41] In fact, the single onset and coda consonants were excerpted from the recordings, and the obstruents were acoustically analyzed. For lack of time, however, no statistical data analyses could be done. This work will wait until after the defense of the dissertation.

vowel sounds. The resemblance between and within learners' perception and production suggests that PA EFL learners share some factors that led to committing the same errors in perception and production. This section discusses the similarities and differences in learners' errors in perception and production and provides pedagogical suggestions to improve the PA EFL learning process at the vowel level.

In light of the weak correlations between each EFL learner's perception and production, the implications will be provided for both levels separately. Carley and Mees (2019) state that "When learning to pronounce a new language, it's essential to get your priorities right. The most important sounds are the ones that can change the meaning of words" (p. 1), referring to the phonemic differences among the sounds in the newly acquired language.

On the perception level, it is important to highlight the vowels' structural features and their frequency as linguistic units and accordingly prioritize them in the learning and/or teaching process, especially in this case during listening tasks. My own data for PA EFL listeners concur with literature data from studies on other L1 Arabic varieties in regard to difficulties in AE vowel perception in general. The most wrongly perceived and problematic vowels for PA EFL learners were the (mid)back vowels, especially, /ɑ/ assimilating to two different PA vowels /u:/ and /a:/; /ɔ/ to PA /a:/ as two parts of a bigger confused vowel cluster that consisted of /u, ɑ, o, ɔ/. The distinction between /ɑ/, and /ɔ/, however, should not be of high priority for PA EFL learners, as native AE speakers generally do not use this distinction. The perception of other vowels of confused hybrid pairs (between tense and lax vowels) or between spectrally nonadjacent vowel pairs as detailed in Chapter 4, which surely affect the perception of the intended speech, should be prioritized. These vowels should be given priority in pedagogical planning and curriculum development by teachers and language educators.

PA EFL learners showed less confusion in AE vowel production. Their most serious issues in PA EFL production of AE were in the back vowel area. Only one confused pair appeared in the front two lax vowels (i.e., /ɪ, ɛ/). In general, better vowel production should follow from a better perception of the AE vowel contrasts, but targeted learning tasks are recommended as well.

7.4.1. Pedagogical implications for curriculum developers

Among the Palestinian TEFL objectives stipulated by the Ministry of Education (MoE), listening and speaking skills are given the same weight as core components of L2 proficiency as the other language skills. However, the implementation of the relevant teaching tasks is far from complete. Several studies highlighted that including these skills in teaching plans and realizing those plans as teaching tasks is an important factor in developing such skills, especially the listening skill, because it is assumed that if such skills will not be assessed, they will not be taught in the first place by the teachers (Richards, 2008: 1). The EFL teaching reality in Palestine is still lagging behind in this sense (Farrah & Halahlah, 2020; Farran et al., 2020; Jabali & Abuzaid, 2017; Rehman et al., 2020). Two immediate changes can be implemented in the Palestinian situation. First, educational supervisors at the school level need to encourage teachers to set up teaching plans that involve explicit and targeted listening and speaking exercises in their daily, mid-term, and/or final exams. Second, the school-leaving exam for senior high school students should also include a formal assessment of listening and speaking skills, and a sufficient weight of marks should be specified for their assessment.

Modern views on assessing these two skills at the learners' cognitive psychology level introduced bottom-up and top-down processing of knowledge. The learners' prior knowledge and comprehension skills acquired by reading aloud written texts are not sufficient to develop learners' listening and speaking skills, which are governed by an interpretive process (Richards, 2008). From this, it follows that written text of conversations and spoken discourses cannot develop learners' abilities in the speaking skill, which require interactive and authentic (native) contexts. Accordingly, these skills will not be activated or developed in learners' cognitive abilities, which require learners to be interacting with interlocutors that make them employ learning mechanisms to acquire, monitor, and evaluate their own perception and production.

7.4.2. Pedagogical implications for teachers

Language educators can provide vowel sound exercises for the L2 vowels that are affected by negative transfer resulting from L1 while including pre-, while-, and after-task interactions and reflections. In this relation, teachers are called upon to first consider the patterns of perception errors found in my results and accordingly plan their teaching exercises to allow learners to acquire the marked sounds in L2 and their rules. The difficulty scale of AE vowels for PA EFL learners can also help teachers identify which vowels should create a lasting learning problem and which require less attention. Accordingly, specific learning (listening and speaking tasks) should be presented to the learners, such as minimal pairs presented in Table 7.1 for the most confused AE vowels for the PA EFL learners.

Table 7.4. Minimal pairs for PA EFL learners targeting some confused AE vowels.

	u	ɑ	o	ɔ
Minimal pairs	boot	bot	boat	bought
	cooler	collar	coaler	caller

Note that the organization of the words in Table 7.1 is set so to distance adjacent vowels from each other. It is suggested that if not the four members of any of these sets are to be learned at once, at least adjacent AE vowels should not follow each other to facilitate the learning process. By doing so, it may be easier for them to spot the larger differences between distant vowels among each confusing cluster sooner than between adjacent ones. Surely, for the production of overlapping vowels, it is important that learners learn how to control their vocal tracts to specify correct muscle movements and positioning of the tongue. For example, to pronounce AE /ɔ/, learners should be instructed to lower the tongue and pull it backward while rounding the lips. Audio-visual aids and/or interactions with native speakers inside or outside the school context can be very helpful.

Other vowel contrasts that include the (mid)front vowels where learners show better perception performance can be better separated using the same technique of minimal pairs as a successful and well-tested technique in L2 acquisition.

It is important to recall here that EFL learners showed good to perfect separation between AE vowels in the tenseness feature. Therefore, only minor follow-up efforts are needed to project this good separation on the newly acquired separated vowels. For example, it should be explicitly mentioned for learners that long vowels preceding voiced consonants are typically longer than long vowels preceding voiceless consonants. Additionally, among the errors that PA EFL learners commit is the over-shortening of all EFL vowels. In this situation, some clear examples and drills should target this issue.

As my results are not applicable in specific school contexts, teachers are advised to determine which vowels their students are struggling with and to provide targeted learning tasks that precisely implement their teaching techniques and lesson plans that help learners overcome these issues. Teachers could start by building on what learners already have in their L1 and can be positively transferred to the L2. This highlights the importance that teachers should be aware of the cross-linguistic similarities and differences between learners' L1 and the L2 they are teaching. Afterward, the teaching of new sounds should be built on different tasks that make the learners aware that these sounds are new, and extra attention should be given to acquiring them. At later stages, suprasegmental features should be introduced for the learners as phonemic/distinctive features of EFL that affect the speaker's intelligibility or the listener's comprehension.

Among the reviewed models, L2LP highlights the importance of input quality in L2 acquisition. With the goals of intelligibility and proficiency as stipulated by the Palestinian MoE for the English language, the L2LP model claims that EFL listeners' mental response can be predicted based on the quality and acoustic properties of the received input. Accordingly, it is suggested that teachers are required to provide a sufficient (large) amount of native or native-like input to allow learners to acquire an optimal L2 perception. This can be better developed by engaging learners in authentic communication tasks in which both their perception and production are activated and instantly assessed. Currently, PA EFL learners do not receive such input or assessment (Farran et al., 2020). Rather, the main source of input is the Arabic-accented EFL teacher, who was also taught by nonnative Arabic-accented EFL teachers, which creates an infinite loop leading to language fossilization. Older studies have stressed the importance of authentic input for L2 learners and the undesirability of nonauthentic input, which compromises L2 learners' perception and leads to "incomplete approximation to L2 phonetic norms" (Flege & Eefting, 1987: 81). In this connection, a new curriculum development, away from the currently implemented one, is suggested to focus on the effects of the PA variety on EFL perception. If not possible, at least remedial tasks should be introduced to the upcoming prints (revisions) as well as extracurricular interactive drills. In this relation and more importantly, the assessment of the learner's perception and production should be reintroduced for both the learners and the teachers to create an atmosphere where learners know the learning aims and the teachers know the teaching aims for these two skills in particular. Previous studies show that these two skills are the least, if at all, assessed in current PA EFL teaching (Farran et al., 2020).

Another problem at the social level is the transliteration method adopted to transfer the exact pronunciation of English words into Arabic but forces them to fit the Arabic prosodic and metric system. This strategy involves overturning L2 new sounds into their similar equivalents in Arabic, e.g., 'television' /ˈtɛ.ləˌvɪ.ʒən/ in AE is pronounced in Arabic as /ta.laˈfi.zu:n/, /telˈfi.zu:n/, or some other nonauthentic pronunciation of the English word. This recurrent strategy will impede learners from receiving a pure and authentic source of input inside or even in extramural activities, which in turn can also lead to L2 fossilization.

7.4.3. Pedagogical implications for PA EFL learners

Many studies have set the goal of achieving intelligible and comprehensible production of speech as an individualized task where each learner experiences certain difficulties that require targeted attention. Moreover, as the learning process is becoming increasingly student-centered, the focus of the teaching process is to brace learners with learning strategies to become autonomous. This shift in the teacher role to become more of a monitor, facilitator, or even as a coach also needs to be parallel with awareness of the students' new roles as autonomous learners. Accordingly, the whole language learning process is better seen as a personalized experience, and learners can monitor their own progress and errors outside the school context, assess their own level of proficiency, and target the less developed features in their L2. At present, learners can make use of the unlimited access to pure L2 input from social media websites and even live communication with native speakers. Following an Informal Digital Learning of English (IDLE), EFL learners can benefit from the available dataset specified for such purposes and available online for EFL practice. In this relation, PA EFL learners are recommended to record themselves speaking in the L2 and then to listen and assess their own production of vowels and other segmental and suprasegmental features.

7.5. Future and further studies

The present dissertation explored the perception and production of AE monophthongs by PA EFL learners. As seen in the literature review and the reviewed studies on L1 Arabic perception and production of EFL vowels, the PA variety as L1 received no attention that covers its segmental and the suprasegmental interference in perceiving or producing EFL sounds. As mentioned earlier, the studies presented in this dissertation tackle only the L2 AE monophthongs in perception and production in relation to the PA L1 vowel inventory and its negative transfer in light of the predictions yielded from the assimilation test. However, more data were obtained during the data collection phase concerning the other segmental and suprasegmental features of AE.

PA participants were recorded producing all English consonants in different word positions and two and three consonant clusters in the word-initial and -final positions. Moreover, they were recorded while producing hypothesized problematic consonants in light of the PA L1 varieties and allophones specified for two PA L1 varieties/dialects (urban vs. rural), as reviewed in Section 2.1.2.2 and Tables 2.4–5, and how these reviewed differences might affect PA EFL pronunciation. Additionally, informants read continuous speech from Aesop’s Fables “The North Wind and the Sun” fable (NWS) and its translation into Arabic twice. The fable can be used to elicit most of the phonemic contrasts of English as a foreign language (except /ʒ/). Some of these features (or aspects) could have been included in the present dissertation but were not because of time constraints. Therefore, research studies tackling L1 PA perception and production of the AE consonants and consonant clusters as well as AE stress patterns can enrich this research area even more and add more to my understanding of the main objective of the study, i.e., revealing the effects of L1 the PA sound system on the perception and production of AE corresponding features.

This thesis also produced some interesting findings that need further examination in the future. For example, a more detailed analysis of the pairwise boundaries between the 11 AE monophthongs in perception and production would allow me to determine the shape of the crossover between neighboring vowels. The results in Chapter 4 show the acceptability regions of all 11 AE monophthongs (the ellipses) that were obtained for L1 and L2 listeners. This has not been done before. In the literature, one normally finds stimulus continua between two (maximally three) sounds but never an exhaustively and systematically sampled vowel space. It should be feasible to break this dissertation results down into a large number of pairwise contrasts between adjacent vowels and from that determine the exact shape of the crossover between neighboring vowels. By doing so, I will find not only that the boundaries in the L2 data are not at the same location in the vowel space as in the L1 data but also that the L2 crossover is shallower (more gradual, i.e., less sharply defined – meaning that the L2 listeners are less certain of their identification) than in the L1 data.

It is my hope that this dissertation (and future research) will contribute to alleviating the lamentable conditions for teaching English in the Palestinian school system and raising it up to a higher level.

REFERENCES

Ababneh, I. (2018). English pronunciation errors made by Saudi students. *European Scientific Journal, ESJ, 14*(2), 244–261. doi: 10.19044/esj.2018.v14n2p244

Abercrombie, D. (1967). *Elements of general phonetics*. Chicago: Edinburg University Press

Abu-Salim, I. M. (1982). Syllable structure and syllabification in Palestinian Arabic. *Studies in the Linguistic Sciences, 12*(1), 103–122. Retrieved from https://www.erudit.org/fr/revues/rql/1986-v16-n1-rql2927/602581ar/

Abu-Salim, I. M. (1986). Vowel shortening in Palestinian Arabic. *Lingua, 68*(2-3), 223–240. doi: 10.1016/0024-3841(86)90005-7

Al-Abdely, A., & Thai Yap, N. (2016a). Learning English vowels by Iraqi EFL learners: Perceived difficulty versus actual performance. *3L The Southeast Asian Journal of English Language Studies, 22*(1), 1–18. doi: 10.17576/3l-2016-2201-01

Al-Abdely, A., & Thai Yap, N. (2016b). The interrelation between the perception and production of English monophthongs by speakers of Iraqi Arabic. *Pertanika Journal of Social Sciences and Humanities, 24*(S), 1-10. Retrieved from: http://psasir.upm.edu.my/id/eprint/29338/

Al-Ani, S. H. (1970). *Arabic phonology*. Paris: de Gruyter.

Al-Badawi, K. (2013). An analysis of phonetic, morphological and syntactic errors in English: A case study of Saudi BA students at King Khalid University. *International Journal of Social Science and Humanity, 2*(6), 536–538. doi: 10.7763/ijssh.2012.V2.165

Al-Badawi, M., & Salim, J. (2014). The perception of English vowels by Arab EFL learners: A case study of university students at Zarqa University. *Research on Humanities and Social Sciences, 4*(20), 23–26. Retrieved from https://iiste.org/Journals/index.php/RHSS/article/view/15784

Almbark, R. (2011). Production and perception of SSBE vowels by Syrian Arabic speakers. In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in SLA of speech: New sounds 2010* (Vol. 2, pp. 15–26): Peter Lang: International Verlag der Wissenschaften.

Almbark, R. (2012). The perception and production of SSBE vowels by Syrian Arabic learners: The foreign language model. (Unpublished doctoral dissertation). University of York, UK.

Almbark, R. (2014). Perception and production of SSBE vowels by foreign language learners: Towards a Foreign Language Model. *Concordia University Working Papers in Applied Linguistics*, 5, 3–18.

Almbark, R., & Hellmuth, S. (2015). Acoustic analysis of the Syrian Arabic vowel system. In: *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow.

Ali, E. M. T. (2011). *Speech intelligibility problems of Sudanese learners of English*. (LOT dissertation series nr. 170). Utrecht: LOT.

Ali, E. M. T. (2013). Pronunciation problems: Acoustic analysis of the English vowels produced by Sudanese learners of English. *International Journal of English and Literature, 4*(10), 495–507.

Al-Mallah, Y. (1983). The emphatic sounds in Classical Arabic: A phonological description. *Bethlehem University Journal, 2*(1), 52–58. Retrieved from https://www.jstor.org/stable/26444494?seq=1#metadata_info_tab_contents

Al-Mazrouei, A., Negm, A., & Kulikov, V. (Submitted). The vowel system of Qatari Arabic: Are short vowels lax and non-peripheral? *Journal of IPA*.

Alqarni, H. (2018). *The production of English vowels by native Arabic speakers.* (MA thesis), Colorado State University, Colorado State University Libraries.

Alshalaan, K. (2020). A comparison between English and Arabic sound systems regarding places of articulation. *Open Access Library Journal, 07*(05), 1–7. doi: 10.4236/oalib.1105679

Al-Wer, E. (2002). Jordanian and Palestinian dialects in contact: vowel raising in Amman. In M. C. Jones & E. Esch (Eds.), *Language Change: The interplay of internal, external and extra-linguistic factors* (pp. 63–80). Berlin, New York: Mouton de Gruyter.

Al-Wer, E. (2007). The formation of the dialect of Amman: From chaos to order. In C. Miller, E. Al-Wer, D. Caubet, & J. Watson (Eds.), *Arabic in the city: Issues in dialect contact and language variation* (pp. 55–76). London: Routledge.

Al-Zoubi, S. (2019). The speech sounds of Arabic language and their effect on learning English pronunciation: A contrastive analysis. *International Journal of Humanities and Social Science, 9*(1), 15–27. doi: 10.30845/ijhss.v9n1p2

Amara, M. H. (2003). Recent foreign language education policies in Palestine. *Language Problems and Language Planning, 27*(3), 217–232. doi: 10.1075/lplp.27.3.02ama

Amir, N., Amir, O., & Rosenhouse, J. (2014). Colloquial Arabic vowels in Israel: A comparative acoustic study of two dialects. *The Journal of the Acoustical Society of America*, *136*(4), 1895–1907. doi: 10.1121/1.4894725

Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 279–291). Cambridge, MA: Newbury House.

Angoujard, J.-P. (1990). *Metrical structure of Arabic*. Dordrecht: Foris.

Astruc, L. (2013). Prosody. In M. J. Jones & R. Knight (Eds.), *The Bloomsbury companion to phonetics* (pp. 126–139). London: Bloomsbury Academic.

Baart, J. L. (2010). *A field manual of acoustic phonetics*. Dallas, TX: SIL International.

Badawi, E.-S. M. (1996). *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*. Cairo: American University in Cairo Press.

Balogné Bérces, K., & Szentgyörgyi, Sz. (2006). *The Pronunciation of English*. Budapest: HEFOP Bölcsész Konzorcium. Retrieved from: http://mek.oszk.hu/04900/04910/

Bergem, D. van (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, *12*, 1–23.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words.* (pp. 167–224). Cambridge, MA: The MIT Press.

Best, C. T. (1995). A direct realist view of crosslanguage speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*(2), 775–794. doi: 10.1121/1.1332378

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O.-S. Bohn & M. Munro (Eds.), *The role of language experience in speech perception and production. A festschrift in honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Bezooijen, R. van, & van Heuven, V. J. (2010). Avant-garde Dutch: An acoustic, perceptual and evaluational study. In: D. R. Preston & N. Niedzielsky (Eds.), *A reader in Sociophonetics* (pp. 357–378). New York: De Gruyter.

Blanc, H. (1953). *Studies in North Palestinian Arabic; Linguistic inquiries among the Druzes of western Galilee and Mt. Carmel*. Jerusalem: Israel Oriental Society.

Blankenship, B. (1991). Vowel perception in a second language. (Unpublished master thesis), University of California, Los Angeles.

Boersma, P. (1998). *Functional phonology.* (Doctoral dissertation), University of Amsterdam. The Hague: Holland Academic Graphics.

Boersma, P., & Heuven, V. J. van (2001). Speak and unSpeak with PRAAT. *Glot International, 5*, 341–347.

Boersma, P., & Weenink, D. (2021). *Praat: Doing Phonetics by Computer*. www.praat.org.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279–304). Baltimore, MD: York Press.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(3), 303–328. doi: 10.1017/s0142716400008912

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*(2), 131–158. doi: 10.1017/s0272263100010792

Bohn, O.-S., & Munro, M. J. (Eds.). (2007). *Language experience in second language speech learning: A festschrift in honor of James Emil Flege*. Amsterdam: John Benjamins.

Broselow, E. (1984). An investigation of transfer in second language phonology. *IRAL: International Review of Applied Linguistics in Language Teaching, 22*(4), 253–269. doi: 10.1515/iral.1984.22.4.253

Brown, A. (2015). Syllable Structure. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 85–105). Chichester: Wiley-Blackwell.

Brown, A., & Oyer, S. (2013). Vowel project: Analysis of a native-Arabic speaker. *Linguistic Portfolios, 2*, 31–46. Retrieved from https://repository.stcloudstate.edu/stcloud_ling/vol2/iss1/4/

Brown, H. D. (1994). *Principles of language learning and teaching* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Buckingham, L. (2014). Attitudes to English teachers' accents in the Gulf. *International Journal of Applied Linguistics, 24*(1), 50–73. doi: 10.1111/ijal.12058

Canepari, L. (2007). *A handbook of pronunciation: English, Italian, French, German, Spanish, Portuguese, Russian, Arabic, Hindi, Chinese, Japanese, Esperanto*. Munich: LINCOM Europa.

Carley, P., & Mees, I. M. (2019). *American English Phonetics and Pronunciation Practice* (1st ed.). London: Routledge. doi: 10.4324/9780429492228

Catford, J. (2002). *A practical introduction to phonetics* (2nd ed.). Oxford: Oxford University Press.

Celce-Murcia, M., Brinton, D.M., Goodwin, J.M. & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. 2nd edition. Cambridge, UK: Cambridge University Press.

Cenoz, J., & Lecumberri, M. L. G. (1999). The acquisition of English pronunciation: learners' views. *International Journal of Applied Linguistics, 9*(1), 3-15. doi: 10.1111/j.1473-4192.1999.tb00157.x

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Cohn, A. C. (2017). Phonology. In M. Aronoff & J. Rees-Miller (Eds.), *The handbook of linguistics* (pp. 185–210). Hoboken, NJ: Wiley-Blackwell.

Colantoni, L., Steele, J., Escudero, P., & Neyra, P. R. E. (2015). *Second language speech: Theory and practice*. Cambridge: Cambridge University Press.

Collins, B., Mees, I. M., & Carley, P. (2019). *Practical English phonetics and phonology*. London, New York: Routledge.

Collins, K. A., & Clément, R. (2012). Language and prejudice: Direct and moderated effects. *Journal of Language and Social Psychology, 31*(4), 376–396. doi: 10.1177/0261927x12446611

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *IRAL: International Review of Applied Linguistics in Language Teaching, 9*(2), 147–160. doi: 10.1515/iral.1971.9.2.147

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Cotter, W. M. (2020). The Arabic dialect of Gaza City. *Journal of the International Phonetic Association*, 1–13. doi: 10.1017/s0025100320000134

Cowell, M. (1964). *A reference grammar of Syrian Arabic*. Washington, DC: Georgetown University Press.

Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.

Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 106–124). Chichester: Wiley-Blackwell.

Dauer, R. (1983). Stress-timing and syllable-timing re-analysed. *Journal of Phonetics*, 11, 51–62.

Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry, 26*(3), 465–498. Retrieved from https://www.jstor.org/stable/4178907

del Rosario Garita, M., González Lutz, M. I. & Solís Pérez, N. (2019). English vowel sounds: Pronunciation issues and student and faculty perceptions. *Actualidades Investigativas en Educación*, 19(3): 33-67. doi: 10.15517/aie.v19i3.38629

Denham, K., & Lobeck, A. (2013). *Linguistics for everyone: An introduction* (2nd ed.). Boston, MA: Wadsworth-Cengage Learning.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315–330. doi: 10.1111/j.1467-1770.1977.tb00124.x

Eckman, F. (1987). The reduction of word-final consonant clusters in interlanguage. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 143–162). Providence, RI: Foris.

Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition, 13*(1), 23–41. doi: 10.1017/S0272263100009700

Eckman, F. R. (1996). A functional-typological approach to second language acquisition theory. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 195–211). San Diego, CA: Academic Press.

Eckman, F. R. (2008). Typological markedness and second language phonology. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 95–115). Philadelphia: John Benjamins.

Eckman, F. R., & Iverson, G. K. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research, 9*(3), 234–252. doi: 10.1177/026765839300900302

Eliot, L. (2013). Single-Sex Education and the Brain. *Sex Roles, 69*(7), 363-381. doi:10.1007/s11199-011-0037-y

Elmahdi, O. E. H., & Khan, W. A. (2015). The Pronunciation Problems Faced by Saudi EFL Learners at Secondary Schools. *Education and Linguistics Research 1*(2), 85–99. doi:10.5296/elr.v1i2.7783

Embarki, M. (2013). Phonetics. In J. Owens (Ed.), *The Oxford handbook of Arabic linguistics* (pp. 23–44). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199764136.013.0002

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization* (LOT Dissertation Series nr 113). Utrecht: LOT. Retrieved from https://www.lotpublications.nl/Documents/113_fulltext.pdf

Escudero, P. (2009). Linguistic perception of “similar” L2 sounds. In P. Boersma & S. S. Hamann (Eds.), *Phonology in perception* (pp. 151–190). Berlin: Mouton de Gruyter.

Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *The Journal of the Acoustical Society of America, 130*(5), EL277–EL283. doi: 10.1121/1.3632043

Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *The Journal of the Acoustical Society of America*, *129*(1), EL1–EL7. doi: 10.1121/1.3525042

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*, 15–31. doi: 10.1016/j.wocn.2018.01.002

Fant, G. (1970). *Acoustic theory of speech production: With calculations based on x-ray studies of Russian articulations*. Berlin, Boston: De Gruyter Mouton.

Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America, 139*(1), EL1–EL5. doi: 10.1121/1.4939608

Farrah, M., & Halahlah, N. (2020). Pronunciation problems among Palestinian English major students in Hebron University. *International Journal of Arabic-English Studies, 20*(1), 203–226. doi: 10.33806/ijaes2000.20.1.11

Farran, B., & Hortobágyi, I. (2020). The linguistic landscape of Ramallah, Palestine: A case study of analysing the dominance of English vs. Arabic as L1 in the Main Street of the city. *SYNERGY - ASE, 16*(2), 242–258. Retrieved from https://www.ceeol.com/search/article-detail?id=886291

Farran, B., Hortobágyi, I., & Gyöngyi, F. (2020). Testing English as a foreign language in Palestine: A case study of INJAZ (GCSE) 2018 English exam. *Hungarian Journal of Applied Linguistics, 20*(1), 1–28. doi: 10.18460/ANY.2020.1.001

Fathi, H., & Qassim, Z. (2020). An acoustic study of the production of Iraqi Arabic vowels. *Journal of Al-Frahedis Arts, 12*(40), 692–704. doi: 10.25130/arts.v12i40.724

Ferguson, C. (1956). The emphatic L in Arabic. *Language*, *32*, 446–452. Retrieved from https://www.jstor.org/stable/410565

Ferrand, C. T. (2018). *Speech science: An integrated approach to theory and clinical practice* (4th ed.). Boston, MA: Pearson.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399–423. doi: 10.2307/3588487

Fischer, W. (1997). Classical Arabic. In R. Hetzron (Ed.), *The Semitic languages* (pp. 187–219). London, New York: Routledge.

Flege, J. E. (1981). The phonological basis of foreign accent. *TESOL Quarterly, 15*(4), 443–455. doi: 10.2307/3586485

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47–65. doi: 10.1016/s0095-4470(19)30537-6

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America, 89*(1), 395–411. doi: 10.1121/1.400473

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217–243). Trier: Wissenschaftlicher Verlag Trier.

Flege, J. E. (2003). *Methods for assessing the perception of vowels in a second language*. Padua: UniPress.

Flege, J. E., & Bohn, O.-S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83): Cambridge University Press.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437–470. doi: 10.1006/jpho.1997.0052

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*(1), 67-83. doi:10.1016/S0095-4470(19)30538-8

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106*(5), 2973–2987.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America, 97*(5), 3125–3134. doi: 10.1121/1.428116

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology, 18*(1), 7–44. doi: 10.1017/s0952675701004006

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *Journal of the Acoustical Society of America*, 90, 1816–1827.

Fowler, C. A. (1981). A relationship between coarticulation and compensatory shortening. *Phonetica*, 38, 35–40.

Fries, C. (1945). *Teaching and learning English as a Foreign language*. Ann Arbor, MI: University of Michigan Press.

Fromkin, V., Rodman, R., & Hyams, N. (2019). *An introduction to language* (11th ed.). Boston, MA: Cengage Learning.

Fry, D. B. (1958). Experiments in the Perception of Stress. *Language and Speech, 1*(2), 126-152. doi: 10.1177/002383095800100207

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*(4), 513–543. doi: 10.1017/s0142716406060383

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*(5), 887–898. doi: 10.1037/0012-1649.28.5.887

Ghazali, S., Hamdi, R., & Barkat, M. (2002). Speech rhythm variation in Arabic dialects. In: Bernard Bel & I. Marlin (eds.), *Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence: Laboratoire Parole et Langage, 127-132.

Ghazali, S., Hamdi, R., & Knis, K. (2007). Intonational and rhythmic patterns across the dialect continuum. In E. Benmamoun (Ed.), *Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic Linguistics Volume XIX: Urbana, Illinois April 2005* (pp. 97–122). Amsterdam: John Benjamins.

Gilakjani, A. P. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education, 1*(1), 1–15. doi: 10.5296/jse.v1i1.924

Gleitman, L., & Wanner, E. (1982). *Language acquisition:* The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language Acquisition: The state of the art* (pp. 3–48). Cambridge: Cambridge University Press.

Gluszek, A., & Dovidio, J. F. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology, 29*(2), 224–234. doi: 10.1177/0261927x09359590

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America, 107*, 2711–2724. doi: 10.1121/1.428657

Gussenhoven, C., & Jacobs, H. (2011). *Understanding phonology*. London: Hodder Education.

Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America, 102*(1), 655–658. doi: 10.1121/1.419712

Hall, N. (2017). Phonetic neutralization in Palestinian Arabic vowel shortening, with implications for lexical organization. *Glossa: A Journal of General Linguistics, 2*(1), 1–23. doi: 10.5334/gjgl.257

Hamdi, R., Ghazali, S., & Barkat-Defradas, M. (2005). Syllable structure in spoken Arabic: A comparative investigation. (pp. 2245‒2248). *Proceedings of Interspeech 2005.* Lisbon. Retrieved from: https://hal.archives-ouvertes.fr/hal-01740924

Harley, H. (2006). *English words: A linguistic introduction*. Malden, MA: Wiley-Blackwell.

Hassan, E. M. I. (2014). Pronunciation problems: A case study of English language students at Sudan University of Science and Technology. *English Language and Literature Studies, 4*(4), 31–42. doi: 10.5539/ells.v4n4p31

Hawks, J. W., & Fourakis, M. S. (1995). The perceptual vowel spaces of American English and Modern Greek: A comparison. *Language and Speech, 38*(3), 237–252. doi: 10.1177/002383099503800302

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28*(4), 377–396. doi: 10.1006/jpho.2000.0121

Heeringa, W., & Van de Velde, H. (2018). *Visible Vowels: A tool for the visualization of vowel variation. In: proceedings of CLARIN Annual Conference,* Pisa. Retrieved from: https://office.clarin.eu/v/CE-2018-1292-CLARIN2018_ConferenceProceedings.pdf.

Hellmuth, S. (2013). Phonology. In J. Owens (Ed.), *The Oxford handbook of Arabic linguistics* (pp. 45–70). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199764136.013.0003

Herin, B. (2017). Do Jordanians really speak like Palestinians? *Journal of Arabic and Islamic Studies, 13*, 99–114. doi: 10.5617/jais.4629

Heuven, V. J. van (2002). *Boven de klanken / Beyond the segments*. Edita, Amsterdam (32 pp + CD-ROM). Retrieved from https://www.let.leidenuniv.nl/ulcl/faculty/vheuven/oratie-eng.htm

Heuven, V. J. van (2016). An acoustic characterization of English vowels produced by American, Dutch, Chinese and Hungarian speakers. *Hungarian Journal of Applied Linguistics, 16*(2), 1–20. doi: 10.18460/ANY.2016.2.004

Heuven, V. J. van (2018). Notes on the phonetics of word and sentence stress: A cross-linguistic (re) view. In H. van der Hulst, J. Heinz, & R. Goedemans (Eds.), *The study of word stress and accent: Theories, methods and data* (pp. 15–59): Cambridge University Press.
doi: 10.1017/9781316683101.002

Heuven, V. J. van (2022). Resolving the prosody paradox. In A. Post da Silveira, P. Arantes, A. Rodrigues da Silva Júnior & G. Gimenez (Eds.) *Prosody and bilingualism*. Letraria, Araraquara. Retrieved from: https://www.researchgate.net/publication/359025608

Heuven, V. J. van, Afshar, N., & Disner, S. (2020). Mapping perceptual vowel spaces in native and foreign language: Persian learners of English compared with American native speakers. In S. Bátyi & Z. Lengyel (Eds.), *Kétnyelvüség: Magyar és nem Magyar kontextus. Tanulmányok Navracsics Judit köszöntésére/Bilingualism: Hungarian and non-Hungarian context. Studies in honor of Judit Navracsics* (pp. 113–130). Veszprém: Pannon Egyetem. Retrieved from: http://hdl.handle.net/1887/3146383

Heuven, V. J. van, & Farran, B. M. M. (2022). The role of duration in the articulation and mental representation of American English vowels by Palestinian Arabic learners of English. Talk presented at the 23rd Summer School of Psycholinguistics, 23-25 May 2022, University of Pannonia, Veszprém, Hungary. Retrieved from: https://www.researchgate.net/publication/361570474

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97*(5), 3099–3111. doi: 10.1121/1.411872

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America, 108*(6), 3013–3022. doi: 10.1121/1.1323463

Hogg, R., & McCully, C. (1987). *Metrical phonology: A coursebook*. Cambridge: Cambridge University Press.

Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.

Holland, C., & Brandenburg, T. (2017). Beyond the front range: The Coloradan vowel space. *The Publication of the American Dialect Society, 102*(1), 9-30. doi: 10.1215/00031283-4295277

Horesh, U. (2020). Palestinian dialects and identities shifting across physical and virtual borders. *Multilingua, 40*(5), 647–673. doi: 10.1515/multi2020-0104

Hosmer, D.W. and Lemeshow, S. (1989) *Applied Logistic Regression*. New York: John Wiley & Sons, Inc.

Hosoda, M., Dietz, J., & Stone-Romero, E. (2010). The effects of foreign accents on employment-related decisions. *Journal of Managerial Psychology, 25*(2), 113-132. doi: 10.1108/02683941011019339

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America, 33*(9), 1174–1178. doi: 10.1121/1.1908941

Hubais, A., & Pillai, S. (2010). An instrumental analysis of English vowels produced by Omanis. *Journal of Modern Languages, 20*(1), 1–18. Retrieved from https://jml.um.edu.my/index.php/JML/article/view/3371

Ingram, J. C. L., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25*(3), 343–370. doi: 10.1006/jpho.1997.0048

Jabali, O., & Abuzaid, Y. (2017). Pronunciation errors committed by Palestinian students at An-Najah National University: An analytical approach. *Arab World English Journal, 8*(2), 119–131. doi: 10.24093/awej/vol8no2.8

Johnson, R., Huron, D., and Collister, L. B. (2014). Music and lyrics interactions and their influence on recognition of sung words: an investigation of word frequency, rhyme, metric stress, vocal timbre, melisma, and repetition priming. *Empirical Musicology Review, 9*(1), 2–20. doi: 10.18061/emr.v9i1.3729

Kalaldeh, R. (2018). Acoustic analysis of Modern Standard Arabic vowels by Jordanian speakers. *International Journal of Arabic-English Studies, 18*, 23–48. Retrieved from http://www.ijaes.net/article/FullText/2?volume=18&issue=1

Kaye, A. S. (2001). Diglossia: The state of the art. *International Journal of the Sociology of Language, 2001*(152), 117–130. doi: 10.1515/ijsl.2001.051

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (Ed.), In *Phonetic Linguistics. Essays in honor of Peter Ladefoged* (pp. 115–132). Indianapolis, IN: Academic Press.

Kenstowicz, M. (1986). Notes on syllable structure in three Arabic dialects. *Revue québécoise de linguistique, 16*(1), 101–127. doi: 10.7202/602581ar

Kenstowicz, M. & Kisseberth, C. (1979). *Generative phonology: Description and theory*. New York: Academic Press.

Khalil, S. (2014). Comparative study of the acoustic vowel space of Egyptian English vowels and general American English vowels. *Linguistic Portfolios, 3*, 76–90. Retrieved from https://repository.stcloudstate.edu/stcloud_ling/vol3/iss1/8/

Kharma, N., & Hajjaj, A. (1989). *Errors of English among Arabic speakers*. London: Alfred Place.

Klecka, W. R., 1981. *Discriminant Analysis*. London: Sage Publications.

Klein, W., Plomp, R., & Pols, L. C. W. (1970). Vowel spectra, vowel spaces, and vowel identification. *The Journal of the Acoustical Society of America, 48*(4B), 999–1009. doi: 10.1121/1.1912239

Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: an auditory explanation. *Journal of Phonetics, 16*(2), 153–169. doi: 10.1016/S0095-4470(19)30480-2

Kochem, T., Ghosh, M., Compton, L., & Cotos, E. (2020). *Oral Communication for Non-Native Speakers of English*. Ames, IA: Iowa University Digital Press. Retrieved from https://dx.doi.org/10.31274/isudp.2020.29

Koffi, E. (2013). The acoustic vowel space of Central Minnesota English: Focus on female vowels. *Linguistic Portfolios, 2*(1), 2–20. Retrieved from: https://repository.stcloudstate.edu/stcloud_ling/vol2/iss1/2/

Koffi, E. (2021). *Relevant acoustic phonetics of L2 English* (1st ed.). Boca Raton, FL: CRC Press.

Koopmans-van Beinum, F. J. (1980). *Vowel contrast reduction*. Amsterdam: Academische Pers.

Lababidi, Z., & Park, H. (2017). Chapter 4. Perceptual mapping between Arabic and English consonants. In H. Ouali (Ed.), *Perspectives on Arabic Linguistics XXIX* (pp. 89–126). Philadelphia/Amsterdam: John Benjamins.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology and sound change*. New York: Mouton de Gruyer.

Ladefoged, P. (1999). American English. In: International Phonetic Association (Ed.), *Handbook of the International Phonetic Association* (pp. 41–44). Cambridge, New York: Cambridge University Press.

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Canada: Thomson, Wadsworth.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29*, 98–104. doi: 10.1121/1.1908694

Ladefoged, P., & Disner, S. (2012). *Vowels and consonants* (3rd ed.). Oxford: Wiley-Blackwell.

Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). Australia: Wadsworth.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Lado, R. (1964). *Language teaching: A scientific approach*. New Delhi: McGraw-Hill.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, *5*(3), 253–264.

Lehiste, I., & Peterson, G. E. (1961). Transitions, Glides, and Diphthongs. *Journal of the Acoustical Society of America, 33*(3), 268-277. doi: 10.1121/1.1908638

Lehman, W. P., & Heffner, R.-M. S. (1940). Notes on the length of vowels (III). *American Speech*, 15(4), 377–380. Retrieved from: https://www.jstor.org/stable/487070

Lehn, W., & Slager, W. R. (1959). A contrastive study of Egyptian Arabic and American English: The segmental phonemes. *Language Learning, 9*(1-2), 25–33. doi: 10.1111/j.1467-1770.1959.tb01126.x

Leussen, J.-W. van, & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, *6*, 1000. doi: 10.3389/fpsyg.2015.01000

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249–336.

Lindblom, B. E. F. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*(11), 1773–1781.

Lindblom B. (1990) Explaining phonetic variation: A sketch of the H&H Theory. In W.J.Hardcastle & A. Marchal (Eds.), Speech production and speech modeling. NATO ASI Series (Series D: Behavioral and Social Sciences, vol 55). Springer, Dordrecht. doi: 10.1007/978-94-009-2037-8_16

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*(04). doi: 10.1017/s0272263199004027

Mackay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics, 27*(2), 157–183. doi: 10.1017/s0142716406060231

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Maddieson, I. (2013a). Consonant inventories. In Dryer, Matthew S. & Haspelmath, Martin (eds.). *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/1

Maddieson, I. (2013b). Vowel quality inventories. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/2

Maddieson, I. (2013c). Presence of Uncommon Consonants. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from https://wals.info/chapter/19

Maddieson, I. (2013d). Absence of Common Consonants. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from https://wals.info/chapter/18

Mair, C. (2005). Recent advances in contrastive linguistics and language typology: the spin-off for language teachers. In D. J. Allerton, C. Tschichold, & J. Wieser (Eds.), *Linguistics, language teaching and language learning* (pp. 21–39). Basel: Schwabe.

Major, R. C. (1990). Phonological universals in second language variation. In H. Burmeister & P. Rounds (Eds.), *Proceedings of the Tenth Meeting of the Second Language Research Forum* (pp. 473–488). Eugene, OR: University of Oregon.

Major, R. C. (1996). Markedness in second language acquisition of consonant clusters. In D. R. Preston & R. Bayley (Eds.), *Variation linguistics and second language acquisition* (pp. 75–96). Amsterdam & Philadelphia: John Benjamins.

Major, R. C. (2014). *Foreign accent: The ontogeny and phylogeny of second language phonology*. New York: Routledge.

McCarthy, J., & Prince, A. (1990). Prosodic morphology and templatic morphology. In M. Eid & J. McCarthy (Eds.), *Perspectives on Arabic linguistics II: Papers from the second annual symposium on Arabic linguistics* (Vol. 1, pp. 1–54). Amsterdam: John Benjamins.

McCarus, E. N. (2011). Modern standard Arabic. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich, & A. Zaborski (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 238–262). Leiden: Brill.

McLoughlin, L. (2009). *Colloquial Arabic (Levantine)* (New ed.). London: Routledge.

Mion, G. (2011). Open questions on stress in Arabic: Some sociophonological interpretations. In A. Avram, A. Focşeneanu, & G. Grigore (Eds.), *A festschrift for Nadia Anghelescu* (pp. 343–356). Bucharest: Editura Universitãţii din Bucureşti.

Mitchell, T. F. (1993). *Pronouncing Arabic, 2*. Oxford: Clarendon Press.

Monahan, P. (2002). *Harmony and metrical structure in Palestinian Arabic*. In: A. Agwuele, & S.-H. Park (Eds.), *Proceedings of the 2002 Texas Linguistics Society*. (pp. 182–198). Austin, TX.: Texas Linguistic Forum.

Morgan, J., & Demuth, K. (Eds.). (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*(1), 39-66. doi: 10.1177/002383099303600103

Murray M. J., Flege J. E., & MacKay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied psycholinguistics 17*(3): 313–334.

Mustafawi, E. (2018). Arabic phonology. In E. Benmamoun & R. Bassiouney (Eds.), *The Routledge handbook of Arabic linguistics* (pp. 11–31). New York: Routledge.

Nemser, W. (1971). *An experimental study of phonological interference in the English of Hungarians*. Bloomington: Indiana University Press.

Nespor, M., & Vogel, I. (2008). *Introduction to 2nd Edition of prosodic phonology*. Dordrecht: Foris. (2nd ed., 2008. Berlin: Mouton de Gruyter).

Newman, D., & Verhoeven, J. (2002). Frequency analysis of Arabic vowels in connected speech. *Antwerp Papers in Linguistics, 100*, 77–86. Retrieved from https://dro.dur.ac.uk/4419/

Nikolova-Simic, A. (2010). L1 Interference in the Perception and Production of English Vowels by Arabic Speakers. (Unpublished doctoral dissertation), Alliant International University, Shirley M. Hufstedler School of Education, San Diego.

Nooteboom, S. G. (1981). Speech rate and segmental perception or the role of words in phoneme identification. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 143–150). Amsterdam: North Holland. doi: 10.1016/S0166-4115(08)60188-0

Owens, J. (Ed.). (2013). *The Oxford handbook of Arabic linguistics* (pp. 45–70). Oxford: Oxford University Press. doi: 0.1093/oxfordhb/9780199764136.001.0001

Pallier, C., Sebastián-Gallés, N., & Colomé, A. (1999). *Phonological representations and repetition priming*. *Proceedings of Eurospeech '99*. Budapest, 1907–1910.
Retrieved from http://cogprints.org/928/3/p035.pdf

Palva, H. (1965). *Lower Galilean Arabic: An analysis of its anaptyctic and prothetic vowels with sample texts*. Helsinki: Studia Orientalia.

Park, H., & de Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics, 36*(4), 704–723. doi: 10.1016/j.wocn.2008.06.002

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*(2), 175–184. doi: 10.1121/1.1906875

Peterson, G. E. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*(6), 693–703. doi: 10.1121/1.1908183

Raphael, L. & Bell-Berti, F. (1975). Tongue Musculature and the Feature of Tension in English Vowels. *Phonetica, 32*(1), 61-73. doi: 10.1159/000259686

Rauber, A. S., Escudero, P., Bion, R. A. H., & Baptista, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian. (pp. 2913–2916). In: *Proceedings of Interspeech 2005*. Lisbon. Retrieved from:
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.422.8294&rep=rep1&type=pdf

Rehman, I., Silpachai, A., Levis, J., Zhao, G., & Gutierrez-Osuna, R. (2020). The English pronunciation of Arabic speakers: A data-driven approach to segmental error identification. *Language Teaching Research*, 1–27. doi: 10.1177/1362168820931888

Richards, J. C. (2008). *Teaching listening and speaking from theory to practice*. New York: Cambridge University Press.

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

Rogers, H. (2000). *The sounds of language: An introduction to phonetics*. New York: Longman.

Ryding, K. C. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Ryding, K. C. (2014). *Arabic: A linguistic introduction*. Cambridge: Cambridge University Press.

Saadah, E. (2011). *The production of Arabic vowels by English L2 learners and heritage speakers of Arabic*. (Doctoral dissertation), University of Illinois. Retrieved from: http://hdl.handle.net/2142/24104

Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of sound systems. *Journal of Phonetics, 25*(3), 255–286. doi: 10.1006/jpho.1997.0043

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*(1-4), 209–232. doi: 10.1515/iral.1972.10.1-4.209

Selinker, L. (1992). *Rediscovering Interlanguage*. New York: Longman.

Shachmon, O. (2017). Ḥalabi Arabic as a contact Dialect in Jerusalem. *Journal of Jewish Languages, 5*(1), 49–80. doi: 10.1163/22134638-12340077

Shahin, K. (2011). Palestinian Arabic. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich, & A. Zaborski (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 526–538). Leiden: Brill.

Strange, W., Bohn, O.-S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America, 115*(4), 1791–1807. doi: 10.1121/1.1687832

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *The Journal of the Acoustical Society of America*, *122*(2), 1111–1129. doi: 10.1121/1.2749716

Stevens, K. N. (2000). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Tarone, E. (1988). *Variation in Interlanguage*. London: Edward Arnold.

Tayeh, Q. Daana, H. and Tahaineh, Y. (2012) Consonant Clusters in Toulkarim Palestinian Dialect. *European Journal of Scientific Research 88*(1), 45–84

Testen, D. (1998). *Parallels in Semitic Linguistics*. Leiden: Brill. doi: 10.1163/9789004348424

Thelwall, R., & Sa'adeddin, M. A. (1999). Illustrations of the IPA: Arabic. In: International Phonetic Association (Ed.), *Handbook of the International Phonetic Association* (pp. 51–54). Cambridge: Cambridge University Press.

Trubetzkoy, N. S., (1939). Grundzüge der Phonologie. *Principles of phonology* (C. Baltaxe, 1969, Trans.). Berkeley: University of California Press.

Tsukada, K. (2009). An acoustic comparison of vowel length contrasts in Arabic, Japanese and Thai: Durational and spectral data. *International Journal of Asian Languages* Processing, 19(4), 127–138. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.602.4739&rep=rep1&type=pdf

Tyler, M. D. (2019). PAM-L2 and phonological category acquisition in the foreign language classroom. In A. M. Nyvad, M. Hejná, A. Højen, A. B. Jespersen, & M. H. Sørensen (Eds.), *A sound approach to language matters: In honor of Ocke-Schwen Bohn* (pp. 607–630). Department of English, School of Communication & Culture: Aarhus University.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of nonnative vowel contrasts. *Phonetica, 71*(1), 4–21. doi: 10.1159/000356237

Vennemann, T. (2011). *Preference laws for syllable structure: And the explanation of sound change with special reference to German, Germanic, Italian, and Latin*. Berlin, Boston: De Gruyter Mouton.

Wang, H., & Heuven, V. J. van (2006). Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. *Linguistics in the Netherlands*, *23*, 237–248.
doi: 10.1075/avt.23.23wan

Wang, H., & Heuven, V. J. van (2014). Is a shared interlanguage beneficial? Mutual intelligibility of American, Dutch and Mandarin speakers of English. In R. van den Doel & L. Rupp (Eds.), *Pronunciation matters. Accents of English in The Netherlands and elsewhere* (pp. 175–194). Amsterdam: VU University Press.
Retrieved from: https://scholarlypublications.universiteitleiden.nl/handle/1887/29645

Wang, H. & Heuven V. J. van (2018). Determining native language background from English vowels and obstruents. *Proceedings of the 2018 IEEE Workshop on Information, Forensics and Security, 11-13 December 2018, Hong Kong*. Retrieved from:
https://scholarlypublications.universiteitleiden.nl/handle/1887/82650

Wang, H. & Heuven V. J. van (2019). Acoustic analysis of English obstruents produced by Chinese, Dutch and American speakers. In Szilvia Bátyi (ed.) *Language – teaching, learning, representation, processing* (Studies in Psycholinguistics 6). Pannon Egyetem MFTK, Veszprém, 141-160.
Retrieved from: https://www.researchgate.net/publication/359857624

Wang, X., & Chen, J. (2019). English speakers' perception of Mandarin consonants: The effect of phonetic distances and L2 experience. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 250–254). Melbourne: Australasian Speech Science and Technology Association. Retrieved from
https://assta.org/proceedings/ICPhS2019/papers/ICPhS_299.pdf

Wang, X., & Chen, J. (2020). The acquisition of Mandarin consonants by English learners: The relationship between perception and production. *Languages*, *5*(2), 1–15. Retrieved from:
http://dx.doi.org/10.3390/languages5020020

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly*, *4*(2), 123–130. doi: 10.2307/3586182

Wardhaugh, R. (1974). *Topics in applied linguistics*. Rowley, MA: Newbury House.

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Watson, J. C. E. (2011). Word stress in Arabic. In M. van Oostendorp, C. J. Ewen, E. V. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*, (Vol. 1, pp. 1–29): John Wiley & Sons.

Wells, J. C. (1982). *Accents of English*. Cambridge, UK: Cambridge University Press.

Yamchi, N. (2006), English Teaching and Training Issues in Palestine. *TESOL Quarterly, 40*(4) 861-865. doi: 10.2307/40264318

Yavaş, M. (2011). *Applied English phonology* (2nd ed.). Oxford: Wiley-Blackwell.

Zsiga, E. (2013). *The sounds of language: An introduction to phonetics and phonology (*Linguistics in the World). Malden, MA & Oxford: Wiley-Blackwell.

Zsiga, E. (2020). *The phonetics/phonology interface*. Edinburgh: Edinburgh University Press.

APPENDICES

Appendix 3.1. Acoustic details of American English vowel tokens used in the PAM test.

The vowel tokens were segmented from the first clear glottal pulse following the [h]-noise until the near-silence of the [d] following the vowel. Mean formant frequencies F1 and F2 were measured (using the Burg algorithm in Praat with three formants in a 0 to 3 KHz frequency band), from the vowel onset until either F1 or F2 showed the beginning of a transition to the following [d]. For the semidiphthongs /e/ and/o/, the mean F1 and F2 were computed for the first 50% of the vowel duration only. The results are shown in Table A1.

Table A1. Stimulus analysis of 22 vowel tokens used in the PAM test. F1, F2 (Hz) and duration (ms) of eleven vowel tokens produced by two male American native speakers.

Vowel	Speaker 1			Speaker 2		
	F1	F2	Dur	F1	F2	Dur
i	300	2378	220	296	2163	184
ɪ	399	1989	113	403	1778	136
e	382	2157	293	436	1890	240
ɛ	506	1887	184	524	1688	128
æ	649	1730	273	667	1637	216
ɑ	713	1150	249	620	992	258
ʌ	595	1154	172	523	1186	150
ɔ	550	967	278	589	873	268
o	442	1047	289	424	1213	205
ʊ	433	1339	209	465	1201	160
u	352	1184	245	353	1212	205

The formants were psychophysically scaled to Bark units so that equal distances in the F1-F2-plane correspond to equal auditory distances in vowel quality using the formula in Traunmüller (1990). The resulting vowel plot is shown in Figure A1.

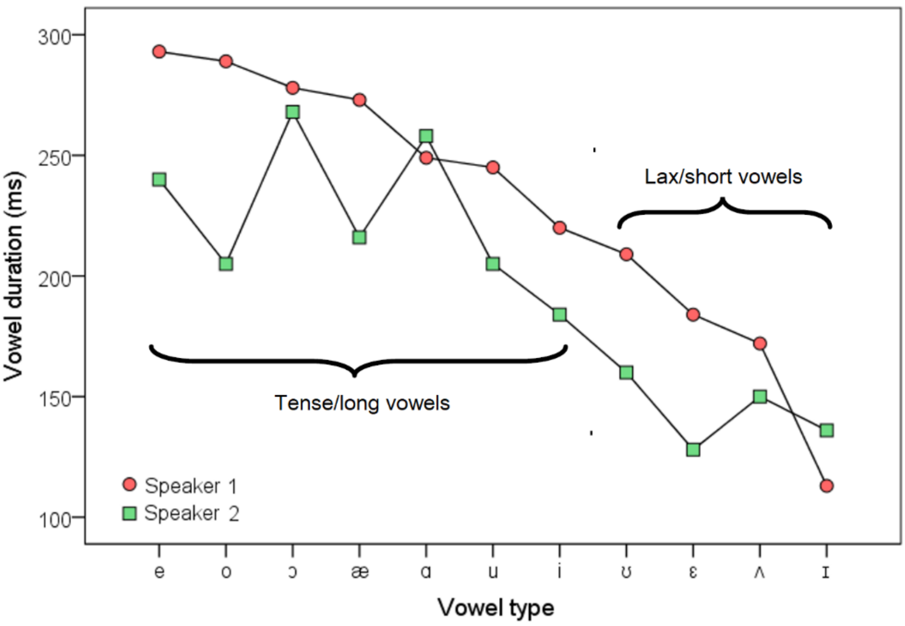

Figure A1. Vowel tokens of Table A1 plotted in the acoustic vowel space defined by F1 (top to bottom, Bark) and F2 (right to left, Bark). Ellipses were drawn by hand and have no theoretical status.

A plot of the vowel durations is shown in Figure A2.

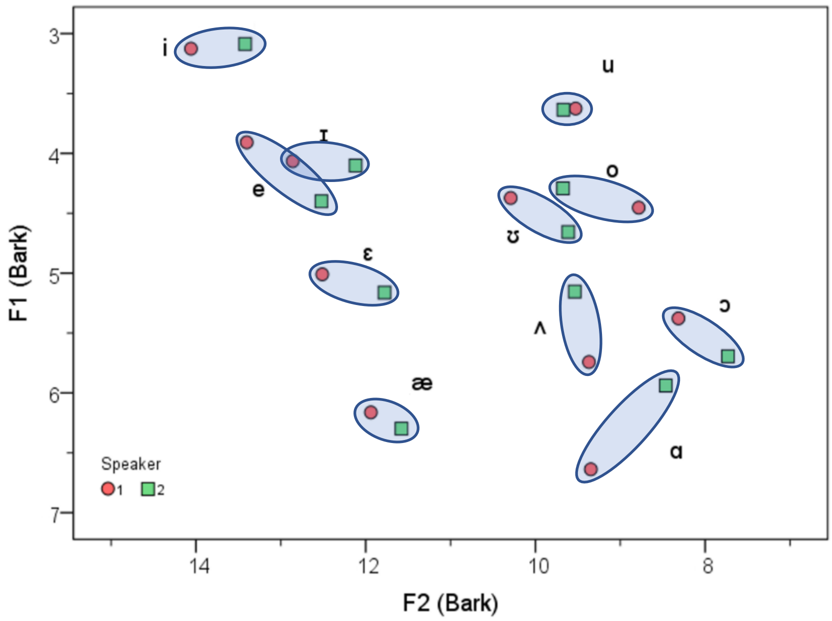

Figure A2. Duration (ms) of 11 American English monophthongs produced by two male native speakers. The vowel types are plotted from left to right in descending order of the duration realized by speaker 1.

There is a clear split in duration between the seven phonetically tense and long vowels and the four lax and short vowels. There are two vowel pairs in Figure A1, the members of which are spectrally close to one another. These are the pairs /ɪ, e/ and /ʊ, o/. These members will nevertheless be distinct by the difference in duration, and by the slight change in quality in the time course of the semi-diphthongs /e/ and /o/.

Appendix 3.2 PA participant biographic data

Code#	Name	Gender	Age	L1	Born in
M01	Nidal	M	16 Years, 5 Months	Arabic	Palestine
M02	Amir	M	16 Years, 8 Months	Arabic	Palestine
M03	Mahdy	M	16 Years, 6 Months	Arabic	Palestine
M04	Moheeb	M	17 Years, 0 Months	Arabic	Palestine
M05	Uday	M	16 Years, 3 Months	Arabic	Palestine
M06	Yanal	M	17 Years, 1 Months	Arabic	Palestine
M07	Noor	M	16 Years, 7 Months	Arabic	Palestine
M08	Mohammad Sh.	M	17 Years, 1 Months	Arabic	Palestine
M09	Braa'	M	16 Years, 2 Months	Arabic	Palestine
M10	Mahmoud	M	16 Years, 1 Months	Arabic	Palestine
M11	Laith	M	16 Years, 10 Months	Arabic	Palestine
M12	Khaled	M	16 Years, 3 Months	Arabic	Palestine
M13	Mohammad H.	M	16 Years, 2 Months	Arabic	Palestine
M14	Adnan	M	16 Years, 3 Months	Arabic	Palestine
M15	Alaa'	M	16 Years, 6 Months	Arabic	Palestine
M16	Sameh	M	17 Years, 1 Months	Arabic	Palestine
M17	Loay	M	16 Years, 6 Months	Arabic	Palestine
M18	Mustafa	M	16 Years, 11 Months	Arabic	Palestine
M19	Abdullah	M	16 Years, 9 Months	Arabic	Palestine
M20	Waseem	M	16 Years, 8 Months	Arabic	Palestine
Code#	Name	Gender	Age	L2	Born in
F01	Islam	F	16 Years, 4 Months	Arabic	Palestine
F02	Lina	F	17 Years, 1 Months	Arabic	Palestine
F03	Noha	F	17 Years, 0 Months	Arabic	Palestine
F04	Nour	F	16 Years, 11 Months	Arabic	Palestine
F05	Roaa	F	16 Years, 3 Months	Arabic	Palestine
F06	Sadil	F	16 Years, 1 Months	Arabic	Palestine
F07	Shatha	F	15 Years, 10 Months	Arabic	Palestine
F08	Sujood	F	16 Years, 11 Months	Arabic	Palestine
F09	Tala	F	16 Years, 8 Months	Arabic	Palestine
F10	Tasnim	F	16 Years, 6 Months	Arabic	Palestine
F11	Reem	F	16 Years, 5 Months	Arabic	Palestine
F12	Batoul	F	16 Years, 1 Months	Arabic	Palestine
F13	Niveleen	F	16 Years, 5 Months	Arabic	Palestine
F14	Haneen	F	17 Years, 1 Months	Arabic	Palestine
F15	Rana	F	16 Years, 5 Months	Arabic	Palestine
F16	Ruba	F	16 Years, 7 Months	Arabic	Palestine
F17	Hala M.	F	16 Years, 2 Months	Arabic	Palestine
F18	Hala H.	F	16 Years, 8 Months	Arabic	Palestine
F19	Malak	F	16 Years, 5 Months	Arabic	Palestine
F20	Sally	F	16 Years, 6 Months	Arabic	Palestine

Age mean = 16;4; Age range 15;10 – 17;1

Appendix 3.3.A. Samples of legal guardian consent form for Palestinian Arabic learners of EFL (Arabic version).

نموذج موافقة ولي الأمر على مشاركة الطالب في تجربة بحثية

العنوان: تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً – دراسة تجريبية.

مقدمة

الهدف من هذا النموذج هو اطلاعك (كولي أمر الطالبـ/ـة) على معلومات قد تؤثر على قرارك بمشاركة الطالبـ/ـة في هذه التجربة البحثية وهي تجربة بحثية عملية تهدف إلى قياس مدى تأثير اللهجات الفلسطينية المختلفة على اكتساب اللغة الإنجليزية وفهمها لدى طلاب المرحلة الثانوية. بعد موافقتك، سيتم في هذه التجربة تسجيل كلمات بصوت الطالبـ/ـة باللغتين العربية والإنجليزية. ستستخدم هذه المواد المسجلة لأغراض البحث العلمي فقط ولن يتم الكشف عن أي من معلومات الطالب الشخصية. اشتراك الطالب في التجربة طوعي كلياً ولن يؤثر على تحصيله الدراسي بأي شكل كان. بإمكانك التواصل مع الباحث ليجيب عن أي أسئلة أخرى لديكم.

البحث

خلال التجربة سيقوم الطالب بالاستماع إلى كلمات وتحديد مدى تشابهها مع اللغة الأخرى وقراءة نص جهرياً وتسمية رسومات وتعبئة استبانة. مجمل التجربة يستغرق 25 دقيقة وتقام داخل حرم المدرسة وبحضور من الهيئة التدريسية. كل معلومات الطلاب الشخصية ستبقى سرية ولن يتم الإفصاح عنها. سيشارك في التجربة حوالي 60 طالباً وطالبة، كلٌ على حِدة.

سيتم تسجيل صوت الطالبـ/ـة لأغراض البحث العلمي فقط ولا يوجد هنالك أي مخاطر تذكر للمشاركة في هذه التجربة. بالإضافة إلى موافقة ولي الأمر، يجب أن يوافق الطلاب على المشاركة في التجربة. إن لم يوافق الطلاب على المشاركة فسيتم استبعادهم من التجربة البحثية ولن يكون هنالك أي عقوبات أو تأثير على تحصيلهم الدراسي. المشاركة طوعية بالكامل ويمكن للمشاركـ/ـة الانسحاب في أي وقت قبل أو خلال التجربة. سيتم إخفاء كل معلومات الطالبـ/ـة الشخصية بشكل كامل. وستحفظ كل السجلات والتسجيلات ولن يصل إليها إلا فريق البحث وسيتم التخلص منها بعد الانتهاء من تحليلها.

نص الموافقة

أنا الموقع أدناه، ولي أمر الطالبـ/ـة: ______________________________ أوافق على مشاركتهـ/ـا في التجربة البحثية المذكورة آنفا وذلك بعد الاطلاع على تفاصيلها، وعليه أوقع.

التوقيع: ______________________________ التاريخ: ___/___/2020

معلومات التواصل:

اسم الباحث لدرجة الدكتوراة: بشار محمد فران
المشرف: بروفيسور فنسنت فان هوفن
جامعة بانونيا، هنغاريا
جوال رقم: 05[redacted]
بريد إلكتروني: bf[redacted]

العدد ___من ___

نموذج موافقة ولي الأمر على مشاركة الطالب في تجربة بحثية

العنوان: حول تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً- دراسة تجريبية.

مقدمة

الهدف من هذا النموذج هو اطلاعك (كولي امر الطالب/ة) على معلومات قد تؤثر على قرارك بمشاركة الطالب/ة في هذه التجربة البحثية وهي تجربة بحثية عملية تهدف الى قياس مدى تأثير اللهجات الفلسطينية المختلفة على اكتساب اللغة الإنجليزية وفهمها لدى طلاب المرحلة الثانوية. بعد موافقتك، سيتم في هذه التجربة تسجيل كلمات بصوت الطالب/ة باللغتين العربية والانجليزية. ستستخدم هذه المواد المسجلة لأغراض البحث العلمي فقط ولن يتم الكشف عن أي من معلومات الطالب الشخصية. اشتراك الطالب في التجربة طوعي كلياً ولن يؤثر على تحصيله الدراسي بأي شكل كان. بإمكانك التواصل مع الباحث ليجيب عن اي أسئلة أخرى لديكم.

البحث

- خلال التجربة سيقوم الطالب بالاستماع الى كلمات وتحديد مدى تشابهها مع اللغة الأخرى وقراءة نص جهرياً وتسمية رسومات وتعبئة استبانة. مجمل التجربة يستغرق 25 دقيقة وتقام داخل حرم المدرسة وبحضور من الهيئة التدريسية. كل معلومات الطلاب الشخصية ستبقى سرية ولن يتم الإفصاح عنها. سيشارك في التجربة حوالي 60 طالب وطالبة، كلٌ على حدة.

سيتم تسجيل صوت الطالب/ة لأغراض البحث العلمي فقط ولا يوجد هنالك أي مخاطر تذكر للمشاركة في هذه التجربة. بالإضافة الى موافقة ولي الامر، يجب ان يوافق الطلاب على المشاركة في التجربة. ان لم يوافق الطلاب على المشاركة فسيتم استبعادهم من التجربة البحثية ولن يكون هنالك أي عقوبات او تأثير على تحصيلهم الدراسي. المشاركة طوعية بالكامل ويمكن للمشارك/ة الانسحاب في أي وقت قبل او خلال التجربة. سيتم إخفاء كل معلومات الطالب/ة الشخصية بشكل كامل. وستحفظ كل السجلات والتسجيلات ولن يصل اليها الا فريق البحث وسيتم التخلص منها بعد الانتهاء من تحليلها.

نص الموافقة

انا الموقع ادناه، ولي أمر الطالب/ة: [redacted] أوافق على مشاركته/ا في التجربة البحثية المذكورة آنفا وذلك بعد الاطلاع على تفاصيلها، وعليه أوقع.

التوقيع: ________ التاريخ: ١ / ٢ /2020

معلومات التواصل:

اسم الباحث: د. بشار محمد فران

المشرف: بروفيسور فنسنت فان هوفن

جامعة بانونيا، هنغاريا

جوال رقم: 05[redacted]

بريد الكتروني: bfa[redacted]

العدد ___ من ___

نموذج موافقة ولي الأمر على مشاركة الطالب في تجربة بحثية

العنوان: حول تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً- دراسة تجريبية.

مقدمة

الهدف من هذا النموذج هو اطلاعك (كولي امر الطالب/ة) على معلومات قد تؤثر على قرارك بمشاركة الطالب/ة في هذه التجربة البحثية وهي تجربة بحثية عملية تهدف الى قياس مدى تأثير اللهجات الفلسطينية المختلفة على اكتساب اللغة الإنجليزية وفهمها لدى طلاب المرحلة الثانوية. بعد موافقتك، سيتم في هذه التجربة تسجيل كلمات بصوت الطالب/ة باللغتين العربية والانجليزية. ستستخدم هذه المواد المسجلة لأغراض البحث العلمي فقط ولن يتم الكشف عن أي من معلومات الطالب الشخصية. اشتراك الطالب في التجربة طوعي كلياً ولن يؤثر على تحصيله الدراسي بأي شكل كان. بإمكانك التواصل مع الباحث ليجيب عن اي أسئلة أخرى لديكم.

البحث

- خلال التجربة سيقوم الطالب بالاستماع الى كلمات وتحديد مدى تشابهها مع اللغة الأخرى وقراءة نص جهرياً وتسمية رسومات وتعبئة استبانة. مجمل التجربة يستغرق 25 دقيقة وتقام داخل حرم المدرسة وبحضور من الهيئة التدريسية. كل معلومات الطلاب الشخصية ستبقى سرية ولن يتم الإفصاح عنها. سيشارك في التجربة حوالي 60 طالب وطالبة، كلّ على حِدة.

سيتم تسجيل صوت الطالب/ة لأغراض البحث العلمي فقط ولا يوجد هنالك أي مخاطر تذكر للمشاركة في هذه التجربة. بالإضافة الى موافقة ولي الامر، يجب ان يوافق الطلاب على المشاركة في التجربة. ان لم يوافق الطلاب على المشاركة فسيتم استبعادهم من التجربة البحثية ولن يكون هنالك أي عقوبات او تأثير على تحصيلهم الدراسي. المشاركة طوعية بالكامل ويمكن للمشارك/ة الانسحاب في أي وقت قبل او خلال التجربة. سيتم إخفاء كل معلومات الطالب/ة الشخصية بشكل كامل. وستحفظ كل السجلات والتسجيلات ولن يصل اليها الا فريق البحث وسيتم التخلص منها بعد الانتهاء من تحليلها.

نص الموافقة

انا الموقع ادناه، ولي أمر الطالب/ة: ______[redacted]______ أوافق على مشاركته/ا في التجربة البحثية المذكورة آنفا وذلك بعد الاطلاع على تفاصيلها، وعليه أوقع.

التوقيع: ______________ التاريخ: 2020/3/1

معلومات التواصل:

اسم الباحث: د. بشار محمد فران

المشرف: بروفيسور فنسنت فان هوفن

جامعة بانونيا، هنغاريا

جوال رقم: 05[redacted]

بريد الكتروني: bfa[redacted]

العدد ___من ___

Appendix 3.3.B. Legal guardian consent form for Palestinian Arabic learners of EFL (English translated version).

Parental Permission for Children Participation in Research

Title: Perception and Production of American English Sounds by Palestinian Arabic Adolescents

Introduction

The purpose of this form is to provide you (as the parent of a prospective research study participant) with information that may affect your decision as to whether or not to let your child participate in this research study. Upon your approval, your child will be asked to participate in a research study about how different Palestinian dialects may affect the perception and the production of English as a foreign language. The purpose of this study is to highlight how certain features of each dialect have positive or negative effects on the perception and the production of EFL.

You can contact the person conducting the research to answer any further questions.

The research

If you allow your child to participate in this study, they will be asked to listen to sounds and decide how much these sounds resemble different language sounds; reading aloud texts; filling in a questionnaire; and naming pictures. This study will take 20 minutes inside the school and will be for one time only with the presence of school staff. There will be 40 or less participants in this study.

Your child's voice will be audio recorded for scientific purposes solely. There are no foreseeable risks to participating in this study. In addition to your permission, your child must agree to participate in the study. If your child does not want to participate, they will not be included in the study and there will be no penalty. The participation is voluntarily and participants can withdraw before or during the experiment. Anonymity of your child's personal data will be preserved. Any audio recordings will be stored securely and only the research team will have access to the recordings.

Consent form:

I, the undersigned, am Parent or Legal Guardian of ________________________ do consent him/her to participate in the research above detailed and thereby I sign.

________________________________ ____/____/2020

Signature of Parent(s) or Legal Guardian Date

Contact information:

Researcher: Bashar Farran
University of Pannonia, Hungary.
Supervisor. Prof. Dr. Vincent van Heuven
Mobile number: 05[redacted]
Email: bfa[redacted]

Count ___ of ___

Appendix 3.4.A. Language background questionnaire (Arabic translated version).

Today's date: _ _/_ _/2020 Count: _ _ _/_ _ _

استبيانه حول تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً – دراسة تجريبية

تهدف هذه الاستبانة إلى معرفة اللغات التي استعملتها طيلة حياتك إلى الآن. اهتمامنا منصب على أصوات اللغة الإنجليزية وكيفية نطقها. إن كان هنالك أي سؤال لا تفضلـ/ين الإجابة عليه، أخبر المراقب وبإمكانك تجاوز السؤال.

صُممت هذه الاستبانة لتظهر الخلفية اللغوية لدى المشاركين فيها من أجل تحديد تاريخ ازدواجية اللغة لديهم.

لا يوجد هنالك إجابة صحيحة أو خاطئة لكل سؤال. لذلك عند الإجابة على الأسئلة، راعي أن تُظهر إجابتك حقيقة توجهاتك ومشاعرك تجاه الموقف المحدد من السؤال.

الاسم:....................................... **النوع الاجتماعي:** ( ) ذكر/ ( ) أنثى **الصف:**..............شعبة:..............

تاريخ الميلاد:_ _/_ _/_ 200 **مكان الولادة:** .... مكان السكن:................. مدينة/ قرية/ مخيم

مكان ولادة الأب: .... مدينة/ قرية/ مخيم

وظيفة الأب: .... تعليم الأب: ابتدائي/ إعدادي/ ثانوي/ جامعي

مكان ولادة الأم: .... مدينة/ قرية/ مخيم

وظيفة الأم: .... تعليم الأم: ابتدائي/ إعدادي/ ثانوي/ جامعي

السفر خارج فلسطين/المكان والسنة

مدة الدروس الخصوصية للغة الإنجليزية (إن وجِدَت): ....ساعة/الأسبوع

أين تستمع إلى الإنجليزية أو تتكلمها خارج المدرسة: ....

هل عشتـ/ـي خارج فلسطين؟ ....

إذا كانت إجابتك نعم،

الدولة: ____________ منذ ______ حتى ____

الدولة: ____________ منذ ______ حتى ____

الدولة: ____________ منذ ______ حتى ____

هل تلقيتـ/ي تدريب خاص بنطق اللغة الإنجليزية؟

من الانترنت: ________________ لمدة ______

من دورات محادثة: ______________ لمدة ______

طرق أخرى: ________________

حول ازدواجية اللغة

ضع علامة X مقابل أفضل كلمة تصف حالتك لكل خيار:

تكرار استعمال اللغة	أبداً	نادراً	قلما	أحيانا	عادةً	غالباً	دائماً
	1	**2**	**3**	**4**	**5**	**6**	**7**
1. في البيت، أتحدث اللغة الإنجليزية.							
2. في البيت، أتحدث اللغة العربية.							
3. في المدرسة، أتحدث اللغة الإنجليزية.							
4. في المدرسة، أتحدث اللغة العربية.							
5. في الحفلات ومع الأصدقاء، أتحدث اللغة الإنجليزية.							
6. في الحفلات ومع الأصدقاء، أتحدث اللغة العربية.							
7. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة الإنجليزية.							
8. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة العربية.							

أنماط الاستعمال	أبداً						دائماً
	1	**2**	**3**	**4**	**5**	**6**	**7**
9. مع والداي، أتحدث اللغة الإنجليزية.							
10. مع أقربائي، أتحدث اللغة الإنجليزية.							
11. مع أصدقائي المقربين، أتحدث اللغة الإنجليزية.							

1

قدرات عامة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
12. لدي قدرات موسيقية في العزف أو الغناء.							
13. لدي قدرة قوية على حفظ طريقة لفظ الكلمات.							
14. لدي قدرة قوية على تقليد اللكنات الأجنبية واللهجات.							

قدرات لغوية خاصة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
15. من السهل لدي أن أتواصل باللغة الإنجليزية من خلال الهاتف.							
16. من السهل لدي أن أتواصل باللغة العربية من خلال الهاتف.							
17. الناس يمدحونني بسبب قدرتي على لفظ الكلمات الإنجليزية.							
18. أتكلم اللغة الإنجليزية بقواعدها الصحيحة.							
19. أتكلم اللغة العربية بقواعدها الصحيحة.							
20. ألفظ كلمات اللغة الإنجليزية بشكل حسن.							
21. ألفظ كلمات اللغة العربية بشكل حسن.							

آراء حول تعلم اللغة الإنجليزية	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
22. من المهم التّحدُث باللغة الإنجليزية بقواعدها الصحيحة.							
23. من المهم التّحدُث باللغة العربية بقواعدها الصحيحة.							
24. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة الإنجليزية.							
25. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة العربية.							
26. من المهم نُطق الكلمات الإنجليزية بطريقة صحيحة.							
27. من المهم نُطق الكلمات العربية بطريقة صحيحة.							
28. أنا أنتبه لطريقة نطق الناس للكلمات والأصوات في اللغة الإنجليزية.							
29. أنا أرغب بتحسين نطقي في اللغة الإنجليزية.							
30. أنا أرغب بتحسين نطقي في اللغة العربية.							
31. لو كنت قادرا على ذلك، لنطقت الإنجليزية كما ينطقونها في الدول الأجنبية.							
32. النطق ليس مهما في اللغة الإنجليزية لأنه لا يؤثر على التواصل في هذه اللغة.							

الدافع لتعلم اللغة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
33. التحدث بالإنجليزية بشكل حسن يساعدني في التواصل بشكل أفضل على مواقع التواصل الاجتماعي.							
34. أحاول أن أُصاحب أكبر عدد ممكن من متحدثي اللغة الإنجليزية كلغتهم الأم.							
35. سيحترمني متحدثو اللغة الإنجليزية (كلغتهم الأم) بشكل أكبر إذا تكلمت اللغة الإنجليزية (من قواعد وكلمات) بشكل حسن.							
36. سيحترمني متحدثو الإنجليزية (كلغتهم الأم) بشكل أكبر إذا نطقت كلمات وأصوات اللغة الإنجليزية بشكل حَسن.							
37. اللغة الإنجليزية مهمة لنجاحي وتفوقي في المدرسة.							

انتهت الاستبانة
شكرا لك!

Today´s date: __/__/2020 Count: ___/____

استبيانه حول تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً – دراسة تجريبية

تهدف هذه الاستبانة الى معرفة اللغات التي استعملتها طيلة حياتك الى الان. اهتمامنا منصب على أصوات اللغة الانجليزية وكيفية نطقها. إن كان هنالك أي سؤال لا تفضل/ين الاجابة عليه، أخبر المراقب وبإمكانك تجاوز السؤال.

صُممت هذه الاستبانة لتظهر الخلفية اللغوية لدى المشاركين فيها من أجل تحديد تاريخ ازدواجية اللغة لديهم.

لا يوجد هنالك إجابة صحيحة او خاطئة لكل سؤال. لذلك عند الإجابة على الأسئلة، راعي ان تُظهر اجابتك حقيقة توجهاتك ومشاعرك تجاه الموقف المحدد من السؤال.

الاسم: [redacted] النوع الاجتماعي: (✓) ذكر / ( ) أنثى الصف: أول ثانوي شعبة: ....

تاريخ الميلاد: 2003/10/8 مكان الولادة: نابلس مكان السكن: نابلس (مدينة)/ قرية/ مخيم

مكان ولادة الاب: نابلس .... (مدينة)/ قرية/ مخيم

وظيفة الاب: .... تعليم الاب: ابتدائي/ إعدادي/ ثانوي/ جامعي

مكان ولادة الام: نابلس .... (مدينة)/ قرية/ مخيم

وظيفة الام: ربة منزل .... تعليم الام: ابتدائي/ إعدادي/ ثانوي/ جامعي

السفر خارج فلسطين/المكان والسنة

مدة الدروس الخصوصية للغة الإنجليزية (إن وجدت): .... ساعة/الأسبوع

اين تسمع الى او تتكلم الإنجليزية خارج المدرسة: في مشاهدة الافلام الكرتونية

هل عشت/ي خارج فلسطين؟ ...لا..............

إذا كانت اجابتك نعم،

الدولة: ____________ منذ________ حتى_____

الدولة: ____________ منذ________ حتى_____

الدولة: ____________ منذ________ حتى_____

هل تلقيت/ي تدريب خاص بنطق اللغة الإنجليزية؟

من الانترنت: ________________ لمدة________

من دورات محادثة: ________________ لمدة________

طرق أخرى: ________________

حول ازدواجية اللغة

ضع علامة X مقابل أفضل كلمة تصف حالتك لكل خيار:

دائما 7	غالبا 6	عادة 5	احيانا 4	قلما 3	نادرا 2	ابدا 1	تكرار استعمال اللغة
					X		1. في البيت، اتحدث اللغة الإنجليزية.
X							2. في البيت، اتحدث اللغة العربية.
					X		3. في المدرسة، اتحدث اللغة الإنجليزية.
X							4. في المدرسة، اتحدث اللغة العربية.
				X			5. في الحفلات ومع الاصدقاء، اتحدث اللغة الإنجليزية.
X							6. في الحفلات ومع الاصدقاء، اتحدث اللغة العربية.
			X				7. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة الإنجليزية.
X							8. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة العربية.

دائما 7	6	5	4	3	2	ابدا 1	أنماط الاستعمال
						X	9. مع والداي، اتحدث اللغة الإنجليزية.
			X				10. مع اقربائي، اتحدث اللغة الإنجليزية.
					X		ع أصدقائي المقربين، اتحدث اللغة الإنجليزية.

1

غير موافق بشدة						موافق بشدة	قدرات عامة
7	6	5	4	3	2	1	
						X	12. لدي قدرات موسيقية في العزف او الغناء.
				X			13. لدي قدرة قوية على حفظ طريقة لفظ الكلمات.
						X	14. لدي قدرة قوية على تقليد اللكنات الأجنبية واللهجات.

غير موافق بشدة						موافق بشدة	قدرات لغوية خاصة
7	6	5	4	3	2	1	
		X					15. من السهل لدي أن اتواصل باللغة الإنجليزية من خلال الهاتف
						X	16. من السهل لدي ان اتواصل باللغة العربية من خلال الهاتف.
			X				17. الناس يمدحوني بسبب قدرتي على لفظ الكلمات الإنجليزية.
			X				18. أتكلم اللغة الإنجليزية بقواعدها الصحيحة.
					X		19. أتكلم اللغة العربية بقواعدها الصحيحة.
				X			20. ألفظ كلمات اللغة الإنجليزية بشكل حسن.
						X	21. ألفظ كلمات اللغة العربية بشكل حسن.

غير موافق بشدة						موافق بشدة	اراء حول تعلم اللغة الانجليزية
7	6	5	4	3	2	1	
						X	22. من المهم التحدث باللغة الإنجليزية بقواعدها الصحيحة.
						X	23. من المهم التحدث باللغة العربية بقواعدها الصحيحة.
						X	24. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة الإنجليزية.
						X	25. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة العربية.
						X	26. من المهم نطق الكلمات الإنجليزية بطريقة صحيحة.
						X	27. من المهم نطق الكلمات العربية بطريقة صحيحة.
			X				28. انا أنتبه لطريقة نطق الناس للكلمات والاصوات في اللغة الإنجليزية.
				X			29. انا ارغب بتحسين نطقي في اللغة الإنجليزية.
						X	30. انا ارغب بتحسين نطقي في اللغة العربية.
			X				31. لو كنت قادرا على ذلك، لنطقت الإنجليزية كما ينطقونها في الدول الأجنبية.
	X						32. النطق ليس مهما في اللغة الإنجليزية لأنه لا يؤثر على التواصل في هذه اللغة.

غير موافق بشدة						موافق بشدة	الدافع لتعلم اللغة
7	6	5	4	3	2	1	
						X	33. التحدث بالإنجليزية بشكل حسن يساعدني في التواصل بشكل أفضل على مواقع التواصل الاجتماعي.
		X					34. أحاول ان اصاحب أكبر عدد ممكن من المتحدثين باللغة الإنجليزية كلغتهم الام.
	X						35. سيحترمني متحدثون اللغة الإنجليزية (كلغتهم الام) بشكل أكبر إذا تكلمت اللغة الإنجليزية (من قواعد وكلمات) بشكل حسن.
				X			36. سيحترمني متحدثون الإنجليزية (كلغتهم الام) بشكل أكبر إذا نطقت كلمات واصوات اللغة الإنجليزية بشكل حسن.
						X	37. اللغة الإنجليزية مهمة لنجاحي وتفوقي في المدرسة.

انتهت الاستبانة

شكرا لك!

Today's date: __/__/2020 Count: ___/___

استبيانه حول تأثير اللهجات الفلسطينية على اكتساب اللغة الإنجليزية نطقاً وسمعاً – دراسة تجريبية

تهدف هذه الاستبانة الى معرفة اللغات التي استعملتها طيلة حياتك الى الان. اهتمامنا منصب على أصوات اللغة الانجليزية وكيفية نطقها. إن كان هنالك أي سؤال لا تفضل/ين الإجابة عليه، أخبر المراقب وبإمكانك تجاوز السؤال.

صُممت هذه الاستبانة لتظهر الخلفية اللغوية لدى المشاركين فيها من أجل تحديد تاريخ ازدواجية اللغة لديهم.

لا يوجد هنالك إجابة صحيحة او خاطئة لكل سؤال. لذلك عند الإجابة على الأسئلة، راعي ان تُظهر اجابتك حقيقة توجهاتك ومشاعرك تجاه الموقف المحدد من السؤال.

الاسم: [redacted] **النوع الاجتماعي:** ( ) ذكر/(X) أنثى الصف: 11 شعبة: ٣

تاريخ الميلاد: 2003/3/31 مكان الولادة: نابلس مكان السكن: مدينة مدينة/ قرية/ مخيم

مكان ولادة الاب: مدينة مدينة/ قرية/ مخيم

وظيفة الاب: حلاق تعليم الاب: ابتدائي/ إعدادي/ (ثانوي)/ جامعي

مكان ولادة الام: مدينة مدينة/ قرية/ مخيم

وظيفة الام: ربة منزل تعليم الام: ابتدائي/ إعدادي/ ثانوي/ (جامعي)

السفر خارج فلسطين/المكان والسنة

مدة الدروس الخصوصية للغة الإنجليزية (إن وجدت): 6 ساعات/الاسبوع ساعة/الأسبوع

اين تستمع الى او تتكلم الإنجليزية خارج المدرسة: الانترنت

هل عشت/ي خارج فلسطين؟ لا

إذا كانت اجابتك نعم،

الدولة: ____________ منذ ________ حتى ____

الدولة: ____________ منذ ________ حتى ____

الدولة: ____________ منذ ________ حتى ____

هل تلقيت/ي تدريب خاص بنطق اللغة الإنجليزية؟

من الإنترنت: ____________ لمدة ______

من دورات محادثة: ____________ لمدة ______

طرق أخرى: ____________

حول ازدواجية اللغة

ضع علامة X مقابل أفضل كلمة تصف حالتك لكل خيار:

تكرار استعمال اللغة	ابداً 1	نادراً 2	قلما 3	احياناً 4	عادةً 5	غالباً 6	دائما 7
1. في البيت، اتحدث اللغة الإنجليزية.		X					
2. في البيت، اتحدث اللغة العربية.							X
3. في المدرسة، اتحدث اللغة الإنجليزية.				X			
4. في المدرسة، اتحدث اللغة العربية.							X
5. في الحفلات ومع الاصدقاء، اتحدث اللغة الإنجليزية.		X					
6. في الحفلات ومع الاصدقاء، اتحدث اللغة العربية.							X
7. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة الإنجليزية.						X	
8. بشكل عام وفي الخمس سنوات الماضية، تكلمت اللغة العربية.							X

أنماط الاستعمال	ابداً 1	2	3	4	5	6	دائماً 7
9. مع والداي، اتحدث اللغة الإنجليزية.	X						
10. مع اقربائي، اتحدث اللغة الإنجليزية.			X				
11. مع أصدقائي المقربين، اتحدث اللغة الإنجليزية.		X					

1

قدرات عامة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
12. لدي قدرات موسيقية في العزف او الغناء.							X
13. لدي قدرة قوية على حفظ طريقة لفظ الكلمات.				X			
14. لدي قدرة قوية على تقليد اللكنات الأجنبية و اللهجات.				X			

قدرات لغوية خاصة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
15. من السهل لدي أن اتواصل باللغة الإنجليزية من خلال الهاتف.			X				
16. من السهل لدي ان اتواصل باللغة العربية من خلال الهاتف.	X						
17. الناس يمدحوني بسبب قدرتي على لفظ الكلمات الإنجليزية.				X			
18. أتكلم اللغة الإنجليزية بقواعدها الصحيحة.				X			
19. أتكلم اللغة العربية بقواعدها الصحيحة.		X					
20. ألفظ كلمات اللغة الإنجليزية بشكل حسن.			X				
21. ألفظ كلمات اللغة العربية بشكل حسن.	X						

اراء حول تعلم اللغة الانجليزية	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
22. من المهم التحدث باللغة الإنجليزية بقواعدها الصحيحة.					X		
23. من المهم التحدث باللغة العربية بقواعدها الصحيحة.							X
24. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة الإنجليزية.	X						
25. استمتع بتعلم الكلمات الجديدة والطرق الجديدة لقول الأشياء باللغة العربية.	X						
26. من المهم نطق الكلمات الإنجليزية بطريقة صحيحة.	X						
27. من المهم نطق الكلمات العربية بطريقة صحيحة.	X						
28. انا أنتبه لطريقة نطق الناس للكلمات والاصوات في اللغة الإنجليزية.			X				
29. انا ارغب بتحسين نطقي في اللغة الإنجليزية.	X						
30. انا ارغب بتحسين نطقي في اللغة العربية.	X						
31. لو كنت قادرا على ذلك، لنطقت الإنجليزية كما ينطقونها فى الدول الأجنبية.		X					
32. النطق ليس مهما في اللغة الإنجليزية لأنه لا يؤثر على التواصل في هذه اللغة.							X

الدافع لتعلم اللغة	موافق بشدة						غير موافق بشدة
	1	2	3	4	5	6	7
33. التحدث بالإنجليزية بشكل حسن يساعدني في التواصل بشكل أفضل على مواقع التواصل الاجتماعي.	X						
34. أحاول ان اصاحب أكبر عدد ممكن من المتحدثين باللغة الإنجليزية كلغتهم الام.	X						
35. سيحترمني متحدثون اللغة الإنجليزية (كلغتهم الام) بشكل أكبر إذا تكلمت اللغة الإنجليزية (من قواعد وكلمات) بشكل حسن.		X					
36. سيحترمني متحدثون الإنجليزية (كلغتهم الام) بشكل أكبر إذا نطقت كلمات واصوات اللغة الإنجليزية بشكل حسن.	X						
37. اللغة الإنجليزية مهمة لنجاحي وتفوقي في المدرسة.	X						

انتهت الاستبانة

شكرا لك!

Appendix 3.4.B. Language background questionnaire (English version).

Today's date: _ _/_ _/2020 Count: _ _ _/_ _ _

Perception and Production of American English sounds by Palestinian Arabic Adolescents Questionnaire

The purpose of this questionnaire is to learn about what languages you have used over the course of your lifetime. Our special interest is in the sounds of language, that is pronunciation. If there are any questions you would prefer NOT to answer, just say so and we will skip that question.

This is a Language Background Questionnaire (LBQ). It has been designed to determine the history of your bilingualism. (Flege, 1995, 1996, 2002)

There are no right or wrong answers to each question. However, as you answer each question, form your answers with regard to your general attitudes and feelings towards the situation given.

Name: …………………………… Gender: ( ) Male / ( ) Female. Grade:………/……….

Date of Birth: _ _/_ _/ 200_ Place of Birth: …………… Place of residency:………

Father job:…………… . Education:………………..

Father's place of birth:___________________ (city/ village/ camp)

Mother job…………….. Education:………………….

Mother's place of birth: ___________________ (city/ village/ camp)

Private English L. classes:……. Hours/week

Living abroad/ Place and years

Outside Palestine, where have you lived (for at least 3 months):

a. _____________________ From_______ to ________

b. _____________________ From_______ to ________

c. _____________________ From_______ to ________

Have you received special instruction in English pronunciation?

a. Audio tapes? Which ones:__________________________#Hours of use_____

b. Conversation Classes? Where:______________________ #semesters________

c. Other:____

About Bilingualism

*Put an X under the best choice that reflect your status for each statement:

A.	**Frequency of language use (occasions, social situations)**	Never	Rarely	Sometimes	Frequently	Usually	Always
1.	At home, I speak *English*						
2.	At home, I speak *Arabic*						
3.	At school, I speak *English*						
4.	At school, I speak *Arabic*						
5.	At parties and with friends I speak *English*						
6.	At parties and with friends I speak *Arabic*						
7.	Overall, in the past five years, I have been speaking *English*						
8.	Overall, in the past five years, I have been speaking *Arabic*						

B.	**Frequency of language use (occasions, social situations)**	Never	Rarely	Sometimes	Frequently	Usually	Always
1.	With my parents, I speak *English*						
2.	With my younger relatives I speak *English*						
3.	With my two or three best friends, I speak *English*						

1

C. **General abilities**		Never	Rarely	Sometimes	Frequently	Usually	Always
1.	I have a *strong* musical ability (instruments, singing)						
2.	I have a *good* memory for the way words are pronounced						
3.	I am good at imitating foreign accents and dialects						

D. **Specific language abilities (Self-evaluation)**		Never	Rarely	Sometimes	Frequently	Usually	Always
1.	It is easy for me to communicate in *English* on the telephone						
2.	It is easy for me to communicate in *Arabic* on the telephone						
3.	People compliment me on my ability to pronounce English						
4.	I speak *English* with correct grammar						
5.	I speak *Arabic* with correct grammar						
6.	I pronounce *English* well						
7.	I pronounce *Arabic* well						

E. **Opinions about language learning**		Never	Rarely	Sometimes	Frequently	Usually	Always
1.	It is important to speak *English* grammatically.						
2.	It is important to speak *Arabic* grammatically.						
3.	I enjoy learning new words and new ways of saying things in *English*						
4.	I enjoy learning new words and new ways of saying things in *Arabic*						
5.	It is important to pronounce *English* well.						
6.	It is important to pronounce *Arabic* well.						
7.	I pay attention to how people pronounce English words and sounds						
8.	I want to improve my pronunciation of *English*						
9.	I want to improve my pronunciation of *Arabic*						
10	If it were possible, I would pronounce *English* like people born in the USA						
11.	It is *not* important to pronounce English well because pronunciation doesn't affect how well you can communicate						

F. **Motivation to learn English**		Never	Rarely	Sometimes	Frequently	Usually	Always
1.	Speaking *English* well will help me to communicate better on social media						
2.	I try to have as many American (*native English*) friends as possible						
3.	English L1 speakers will respect me more if I speak (grammar, vocabulary) English well						
4.	English L1 speakers will respect me more if I pronounce *English* well						
5.	English is important for my success at school						

The End.
Thank you!

Researcher: Bashar M. Farran
Affiliation: PhD candidate, University of Pannonia

2

Appendix 3.5. Praat script for the PAM test.

```
"ooTextFile"
"ExperimentMFC 7"
blankWhilePlaying? <no>
stimuliAreSounds? <yes>
stimulusFileNameHead = "Sounds/"
stimulusFileNameTail = ".wav"
stimulusCarrierBefore = ""
stimulusCarrierAfter = ""
stimulusInitialSilenceDuration = 1.5 seconds
stimulusMedialSilenceDuration = 0
stimulusFinalSilenceDuration = 0.5 seconds
numberOfDifferentStimuli = 22
"had_01" ""
"hawed_01" ""
"hayed_01" ""
"head_01" ""
"heed_01" ""
"hid_01" ""
"hod_01" ""
"hoed_01" ""
"hood_01" ""
"hud_01" ""
"whod_01" ""
"had_11" ""
"hawed_11" ""
"hayed_11" ""
"head_11" ""
"heed_11" ""
"hid_11" ""
"hod_11" ""
"hoed_11" ""
"hood_11" ""
"hud_11" ""
"whod_11" ""
numberOfReplicationsPerStimulus = 2
breakAfterEvery = 0
randomize = <PermuteBalancedNoDoublets>
startText = " هذه تجربة سماعية.

بعد سماع الصوت، اختر حرف العلة الأقرب لما سمعته

انقر للبدء "
runText = " اختر حرف العلة الاقرب للصوت الذي سمعت "
pauseText = "يمكنك اخذ استراحة إن أردت، أنقر للمتابعة."
endText = " انتهت التجربة!"
maximumNumberOfReplays = 0
replayButton = 0 0 0 0 "" ""
okButton = 0 0 0 0 "" ""
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0 0
numberOfDifferentResponses = 6
0.2 0.4 0.8 0.9 "40 سُر" "" "u"
0.2 0.4 0.6 0.7 "40 سِر" "" "i"
0.2 0.4 0.4 0.5 "40 سَر" "" "a"
0.6 0.8 0.8 0.9 "40 سور" "" "u:"
0.6 0.8 0.6 0.7 "40 سير" "" "i:"
0.6 0.8 0.4 0.5 "40 سار" "" "a:"
numberOfGoodnessCategories = 5
```

0.25 0.35 0.10 0.20 "1 24 "(سيء) ""
0.35 0.45 0.10 0.20 "2" 24 ""
0.45 0.55 0.10 0.20 "3" 24 ""
0.55 0.65 0.10 0.20 "4" 24 ""
0.65 0.75 0.10 0.20 "5 24 "(جيد) ""

Appendix 4.1 AE native listeners' biographic data

Code#	Name	Gender	Age	L1	Born	Raised
101	Bob1220	M	78	English	Chicago IL	Chicago IL
102	Cathy1220	F	68	English	Culver City	Culver City CA
103	Elaine1220	F	76	English	Cleveland	Cleveland, OH
105	Phil1220	M	60	English	San Francisco	San Francisco CA
110	Anjalee122	F	19	English	Palo Alto CA	San Jose CA
111	Carson122	M	21	English	Connecticut	New Jersey
113	Mac121	M	21	English	Wyoming	Wyoming
121	Gabriel123	M	20	English	Florida	various US states
122	Jaron123	M	19	English	California	California
123	Jordan123	F	21	English	New Jersey	New Mexico
126	Aamani128	F	19	English	California	California
127	Aamuro129	M	18	English	Colorado	Colorado
128	Andrew128	M	21	English	Washington	Washington DC
129	Charli128	M	21	English	Washington	Washington DC
130	Daniel128	F	21	English	Virginia	Zurich
131	Jenna129	M	22	English	USA	USA
132	Sawyer129	M	20	English	North CA	North CA
133	Selden129	M	21	English	New York	New York
134	Sophia129	F	19	English	South CA	South CA
135	Tim129	M	22	English	South CA	Calgary, Canada

Appendix 4.2. Praat script for the perceptual mapping of the AE vowel space test by PA learners of EFL

```
"ooTextFile"
"ExperimentMFC 7"
blankWhilePlaying? <no>
stimuliAreSounds? <yes>
stimulusFileNameHead = "Sounds/"
stimulusFileNameTail = ".wav"
stimulusCarrierBefore = ""
stimulusCarrierAfter = ""
stimulusInitialSilenceDuration = 1.5 seconds
stimulusMedialSilenceDuration = 0
stimulusFinalSilenceDuration = 0.5 seconds
numberOfDifferentStimuli = 86

"V1_1_200" ""
"V1_2_200" ""
"V1_3_200" ""
"V1_4_200" ""
"V1_5_200" ""
"V1_6_200" ""
"V1_7_200" ""
"V1_8_200" ""
"V1_9_200" ""
"V2_1_200" ""
"V2_2_200" ""
"V2_3_200" ""
"V2_4_200" ""
"V2_5_200" ""
"V2_6_200" ""
"V2_7_200" ""
"V2_8_200" ""
"V2_9_200" ""
"V3_2_200" ""
"V3_3_200" ""
"V3_4_200" ""
"V3_5_200" ""
"V3_6_200" ""
"V3_7_200" ""
"V3_8_200" ""
"V4_2_200" ""
"V4_3_200" ""
"V4_4_200" ""
"V4_5_200" ""
"V4_6_200" ""
"V4_7_200" ""
"V4_8_200" ""
"V5_3_200" ""
"V5_4_200" ""
"V5_5_200" ""
"V5_6_200" ""
"V5_7_200" ""
"V6_4_200" ""
"V6_5_200" ""
"V6_6_200" ""
"V7_4_200" ""
"V7_5_200" ""
"V7_6_200" ""
"V1_1_300" ""
```

```
"V1_2_300" ""
"V1_3_300" ""
"V1_4_300" ""
"V1_5_300" ""
"V1_6_300" ""
"V1_7_300" ""
"V1_8_300" ""
"V1_9_300" ""
"V2_1_300" ""
"V2_2_300" ""
"V2_3_300" ""
"V2_4_300" ""
"V2_5_300" ""
"V2_6_300" ""
"V2_7_300" ""
"V2_8_300" ""
"V2_9_300" ""
"V3_2_300" ""
"V3_3_300" ""
"V3_4_300" ""
"V3_5_300" ""
"V3_6_300" ""
"V3_7_300" ""
"V3_8_300" ""
"V4_2_300" ""
"V4_3_300" ""
"V4_4_300" ""
"V4_5_300" ""
"V4_6_300" ""
"V4_7_300" ""
"V4_8_300" ""
"V5_3_300" ""
"V5_4_300" ""
"V5_5_300" ""
"V5_6_300" ""
"V5_7_300" ""
"V6_4_300" ""
"V6_5_300" ""
"V6_6_300" ""
"V7_4_300" ""
"V7_5_300" ""
"V7_6_300" ""

numberOfReplicationsPerStimulus = 1
breakAfterEvery = 0
randomize = <PermuteBalancedNoDoublets>
startText = " .سماعية تجربة هذه

سمعته لما الأقرب العلة حرف اختر ،الصوت سماع بعد

" للبدء انقر
runText = " سمعت الذي للصوت الاقرب العلة حرف اختر"
pauseText = ".للمتابعة أنقر ،أردت إن استراحة اخذ يمكنك"
endText = "!التجربة انتهت"
maximumNumberOfReplays = 0
replayButton = 0 0 0 0 "" ""
okButton = 0 0 0 0 "" ""
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0 0
```

```
numberOfDifferentResponses = 15
0.15 0.30 0.8 0.9 "feel"  40 "" "i"
0.40 0.55 0.8 0.9 ""      40 "" ""
0.65 0.80 0.8 0.9 "fool"  40 "" "u"
0.15 0.30 0.65 0.75 "fill"  40 "" "ɪ"
0.40 0.55 0.65 0.75 ""      40 "" ""
0.65 0.80 0.65 0.75 "full"  40 "" "ʊ"
0.15 0.30 0.5 0.6 "sale"  40 "" "e"
0.40 0.55 0.5 0.6 ""      40 "" ""
0.65 0.80 0.5 0.6 "whole" 40 "" "o"
0.15 0.30 0.35 0.45 "tell"  40 "" "ɛ"
0.40 0.55 0.35 0.45 "null"  40 "" "ʌ"
0.65 0.80 0.35 0.45 "call"  40 "" "ɔ"
0.15 0.30 0.2 0.3 "shall" 40 "" "æ"
0.40 0.55 0.2 0.3 ""      40 "" ""
0.65 0.80 0.2 0.3 "doll"  40 "" "ɑ"

numberOfGoodnessCategories = 3
0.25 0.40 0.05 0.15 "24 " 1 (سيء) ""
0.40 0.55 0.05 0.15 "2" 24 ""
0.55 0.70 0.05 0.15 "3 24 "(جيد) ""
```

Appendix 4.3.A. Results of American Native Listeners of the Perceptual Mapping of the AE Vowel Space Test.

Spectrum		Vowel duration = 200 ms											Vowel duration = 300 ms										
F1	F2	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
237	2357	**19**			1								**19**			1							
237	2031	**19**	1										**19**									1	
237	1746	2	6		1		1	2				**8**	4	3		2			1	1		2	**7**
237	1497	1	1								6	**12**							3		1	4	**12**
237	1278	1			1			2		1	1	**14**	1						1			3	**15**
237	1086	1									1	**18**							1			2	**17**
237	915										2	**18**											**20**
237	764									2	1	**17**									1	1	**18**
237	628										2	**18**										1	**19**
339	2357	**16**	3	1									**19**		1								
339	2031	**10**	9									1	**12**	5	1	2							
339	1746	3	**7**			1					4	5	4	**7**		1			1		1	3	3
339	1497		**9**								5	6	1	**6**			2		4			2	5
339	1278	2						1			6	**11**										7	**13**
339	1086										6	**14**							1			1	**18**
339	915							1			4	**15**									1	3	**16**
339	764											**20**										1	**19**
339	628										2	**18**						1			1	2	**16**
447	2031		**10**		10								1	6	2	**11**							
447	1746		**12**		6	1						1	1	6	1	**10**				1		1	
447	1497		5	1	2			4			**5**	3		4	2	3			4			**4**	3
447	1278		2		3		1	3	1		**8**	2		2		1			5			**11**	1
447	1086							4			**11**	5			1		1		6		1	**11**	
447	915						1	3		**9**	6	1						1	5		**9**	4	1
447	764							1	3	**11**	5							1	2	3	**13**	1	
565	2031		3	2	**15**									1	3	**15**						1	
565	1746		1	1	**18**										4	**15**						1	
565	1497				**15**		1	4						1	3	**9**	2			1		4	
565	1278				3		1	**7**	1	2	5	1		1		1	2	2	**9**			5	
565	1086				1		3	**7**		2	7						1	2	**7**	2	1	7	
565	915						4	**6**	3	4	3							8	1	3	**8**		
565	764						**7**	2	3	6	1	1					1	**8**	2	5	4		
694	1746			4	**12**	3			1						2	7	**9**			2			
694	1497			1	**12**	3		1	3				1		3	3	**11**		1	1			
694	1278				2	2	1	**9**		1	5			1	2	1	**6**	2	3	5			
694	1086						5	**10**	1	1	3						1	6	1	**11**	1		
694	915					3	6	1	**9**		1						3	**10**		5	2		
838	1497			4	3	**13**									2		**15**			3			
838	1278					**9**	3	1	6	1					1		**10**	7		2			
838	1086					7	3		**9**	1							**10**	3		7			
998	1497			1		**16**			3						2		**12**			6			
998	1278		1	2		**10**	3		4						2		**9**	4		5			
998	1086			2		5	4		**8**	1					1		**9**	4		6			
All		**74**	**70**	**19**	**105**	**73**	**44**	**69**	**55**	**42**	**100**	**209**	**82**	**43**	**33**	**82**	**104**	**59**	**58**	**69**	**44**	**83**	**203**

Table A4.1. Number of responses given to 86 synthesized vowel stimuli by 20 native listeners of American English broken down by vowel duration (short = 200 ms, long = 300 ms) and by formant frequencies (F1 and F2 in Hz). The eleven response categories are listed in the second row. The majority response per stimulus vowel is in bold print in a green-shaded cell. Tied votes are indicated in yellow-shaded cells. The casting vote then goes to the response category that maximizes contiguous areas in the vowel space.

Appendix 4.3.B. Results of PA EFL Listeners of the Perceptual Mapping of the AE Vowel Space Test.

Table A4.2. Number of responses given to 86 synthesized vowel stimuli by 41 Palestinian Arabic learners of American English broken down by vowel duration (short = 200 ms, long = 300 ms) and by formant frequencies (F1 and F2 in Hz).

Spectrum		Vowel duration = 200 ms											Vowel duration = 300 ms										
F1	F2	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
237	2357	16	**18**		6			1					**39**		1	1							
237	2031	7	**27**		6					1			**34**	1	1	1	1	1			1	1	
237	1746	3	9		6		4	1	1		**12**	5	7	2	1	1				1	2	2	**25**
237	1497	1	2	2	3		8			1	**14**	10			1	1		1	1	1	9		**27**
237	1278					1	6	2		5	**15**	12		3	1			1	1		2	1	**32**
237	1086						5	2	1	4	12	**17**									10		**31**
237	915						6			6	**15**	14						1	1		1	1	**37**
237	764						4			6	11	**20**		1							6		**34**
237	628						3	1		1	14	**22**							1		5		**35**
339	2357	8	**26**		7								**35**	2	1	3							
339	2031	1	**29**		9						2		**29**	2	4	4		1	1				
339	1746		**16**		4		6	3		1	9	2	5	3	2	5	2	2	1		7	2	**12**
339	1497		5	1	1		6		1	5	**19**	3			1	1		2	1	2	6	4	**24**
339	1278		1	2			8	1	2	3	**17**	7					1	1		2	8	3	**26**
339	1086						11	1	1	7	**15**	6							1		6	2	**32**
339	915		1			1	13	1		2	**15**	8						2	1		5	2	**31**
339	764						12			4	**17**	8			1			2	1	2	10		**25**
339	628						15	1	1	4	9	**11**		1	1						5		**34**
447	2031	1	**27**	1	11	1							3	6	**19**	12				1			
447	1746	1	**33**	1	2	1	1				1	1	2	6	**19**	10		1		2			1
447	1497	1	5	1	5		14		2	1	**11**	1	1	1	9	4		3		6	**13**	2	2
447	1278						17	2	2	1	**17**	2		1	3	1		4		5	**16**	3	8
447	1086		1				19	3		2	**15**	1			2			3		2	**21**	1	12
447	915				1	1	15		2	5	**17**						1	3		6	**21**		10
447	764				1		15	1	3	6	**12**	3		1				3		2	**24**	2	9
565	2031		**27**	4	9			1					4	6	**20**	7				1	3		
565	1746		**28**	1	8	1		2			1		3	3	**24**	6	2		1	2			
565	1497	1	**14**	3	1	4	6	3	1		8			2	**17**	3	2		1	7	5	3	1
565	1278		2	1	1	3	**12**	6	1	2	11	2	1		3		3	4	2	10	**14**	1	3
565	1086				1	2	**16**	5	5		11	1			2		2	3	5	5	**14**	2	8
565	915				1	1	**22**	5		2	10				1			4		3	**24**	5	4
565	764		1				**18**	3	5	1	13				1		1	6		4	**19**	4	6
694	1746		**19**	2	11	5		3			1			1	**26**	4	7		1	1	1		
694	1497	1	5		4	12	3	**14**			2			1	12		**17**		6	2	2		1
694	1278			1	1	6	1	**22**	2	2	6				7	1	9	1	**15**	5	2	1	
694	1086		1			7	6	**18**	2		6	1			7		7	1	**13**	7	5	1	
694	915			1	1	2	**15**	8	2		12				5		2	4	6	4	**12**	2	6
838	1497		2		1	16		**17**	1	1	3				8		**18**		9	5	1		
838	1278		2	1		9		**25**	1		3				10		11		**16**	3		1	
838	1086					4	3	**31**		1	2				7		11	1	**18**	1	1	1	1
998	1497			2		14	1	**22**	1		1				10		9		**16**	3	2		1
998	1278		1	1		6		**29**	2		2				5		15		**18**	1		1	1
998	1086			3		7	1	**28**			2			1	6		9	2	**18**	5			
All		**41**	**302**	**28**	**101**	**104**	**292**	**262**	**39**	**74**	**363**	**157**	**163**	**44**	**238**	**65**	**130**	**57**	**155**	**101**	**283**	**48**	**479**

Appendix 5.1. List of stimulus words in common key words (A) and /hV(r)d/ carrier (B).

	A	B
1.	Now say **need** again.	Now say **heed** again.
2.	Now say **kid** again.	Now say **hid** again.
3.	Now say **played** again.	Now say **hayed** again.
4.	Now say **bed** again.	Now say **head** again.
5.	Now say **bad** again.	Now say **had** again.
6.	Now say **rude** again.	Now say **who'd** again.
7.	Now say **good** again.	Now say **hood** again.
8.	Now say **road** again.	Now say **hoed** again.
9.	Now say **sawed** again.	Now say **hawed** again.
10.	Now say **god** again.	Now say **hod** again.
11.	Now say **card** again.	Now say **hard** again.
12.	Now say **mud** again.	Now say **hud** again.
13.	Now say **word** again.	Now say **heard** again.
14.	Now say **slide** again.	Now say **hide** again.
15.	Now say **employed** again.	Now say **hoyed** again.
16.	Now say **loud** again.	Now say **how'd** again.
17.	Now say **beard** again.	Now say **here'd** again.
18.	Now say **toured** again.	Now say **hoored** again.
19.	Now say **shared** again.	Now say **haired** again.
20.	Now say **bed** again.	Now say **head** again.

Appendix 5.2. Praat Script: Concatenator, tiers creator, and labeler.

```
# Concatenate sound files, add 4 tiers (labeled) & silences boundaries checking
#
# By: Bashar M. M. Farran, 19-July-2020, for his dissertation titled: Perception and
Production of American English sounds by Palestinian Arabic Adolescents
# Select the files to be concatenated and run the script
# Script will pause to check silence boundaries
selectObject: 1
plusObject: 2
plusObject: 3
plusObject: 4
plusObject: 5
plusObject: 6
plusObject: 7
plusObject: 8
plusObject: 9
plusObject: 10
plusObject: 11
plusObject: 12
Concatenate recoverably
	selectObject: "TextGrid chain"
	Extract one tier: 1
	selectObject: "TextGrid chain"
Remove
	selectObject: "Sound chain"
	To TextGrid (silences): 100, 0, -35, 0.1, 0.1, "", "item"
pause Check Silences!
	selectObject: "TextGrid labels"
	plusObject: "TextGrid chain"
Merge
	selectObject: "TextGrid labels"
	plusObject: "TextGrid chain"
Remove
	selectObject: "TextGrid merged"
Insert interval tier: 3, "Segments"
	selectObject: "TextGrid merged"
Insert interval tier: 4, "Formants"
	selectObject: "Sound chain"

print "Script Finished"
```

Appendix 5.3. Praat Script: Extract Formants (by J. J. A. Pacilly)

```
# Extract Formant data using analysis settings in intervals in a TextGrid
#
# J J A Pacilly, 1-July-2021, for Vincent van Heuven
#
# When you run this script, a form will appear on screen. Fill in the name of the sound file
# that contains the vowels from which you want to extract formants and duration.
# In the same folder there should be an accompanying TextGrid with exactly the same name.
# Only the extension must be different. Sound file and Textgrid must be in the same folder as this script.
#
# The script presupposes a TextGrid with four tiers. Tier 1 (top tier) lists the sections of the recording
# Tier 2 (second from top) specifies where items (separated by silences) begin and end. Tier 3 marks of the
# beginning and end of the vowel from which you will extract the formants. The bottom tier has the same
# boundaries as Tier 3 but specifies how many formants should be extracted and the upper limit of the
# frequency band within which formants should be found. This label should contain one (and only one) space,
# which separates the number of formants from the max frequency. The script will fail if this condition
# is not met.
#
# The results are written to a text file called "ExtractedFormants.txt". This output file can be found
# in the same folder as the input files. The results of subsequent extractions (with other input files)
# will be appended to the existing output file.
#
# The output specifies the name of the input file (typically the name of the speaker). Next it lists the
# label of the vowel or word that contains the target vowel. Then follows the vowel duration in milli-
# seconds, and the formant frequencies (in hertz and/or Bark units, depending on your choice).
#
# Formant frequencies are extracted by the Burg algorithm, using the number of formants and frequency
# band as specified in the TextGrid, and default values for other parameters: Time step is automatic,
# Window length = 25 ms, Pre-emphasis from 50 Hz. The output value is the median formant frequency
# measured for the entire vowel segment.

form Extract Formant values
  sentence Filename soundfile.wav
  positive tier 3
  word Label_(regex) ^.+$
  boolean  Convert_to_Unicode no
  sentence Result ExtractedFormants.txt
endform

idSnd = Read from file: filename$
oname$ = selected$("Sound")
idTG = Read from file: oname$ + ".TextGrid"
nrIntervals = Get number of intervals: tier
if convert_to_Unicode
  Convert to Unicode
endif

for interval to nrIntervals
  selectObject: idTG
  vowel$ = Get label of interval: tier, interval
  if index_regex(vowel$, label$)
    ts = Get starting point: tier, interval
    te = Get end point: tier, interval
    duration = (te - ts) * 1000

    appendFile: result$, oname$, tab$, vowel$, tab$, fixed$(duration, 0)

    index3 = Get interval at time: 4, (ts + te) / 2
    fmtSetting$ = Get label of interval: 4, index3
```

```
    maxNrFmt = extractNumber(fmtSetting$, "")
    maxFreq = extractNumber(fmtSetting$, " ")

    selectObject: idSnd
    idSndTmp = Extract part: ts-0.1, te+0.1, "rectangular", 1, "yes"

    idFmt = noprogress To Formant (burg): 0, maxNrFmt, maxFreq, 0.025, 50

    for formant to maxNrFmt
#     freq = Get quantile: formant, ts, te, "Bark", 0.5
      freq = Get quantile: formant, ts, te, "Hertz", 0.5
      appendFile: result$, tab$, fixed$(freq, 0)
    endfor
    appendFileLine: result$, ""

     removeObject: idSndTmp, idFmt
   endif
 endfor

removeObject: idSnd, idTG
```

Appendix 5.4. Descriptive statistics of vowel production data

Mean, standard deviation, range, minimum and maximum for F1 (Hz), F2 (Hz) and duration (ms) of American English vowels produced by Palestinian Arabic EFL learners, aggregated and broken down by gender of speaker.

PA Participants			American English Monophthong										
Gender	Statistic	Parm	i	ɪ	e	ɛ	æ	ɑ	ɔ	o	ʊ	u	ʌ
Male N = 20	Mean	F1	364	442	453	456	666	530	529	481	411	399	614
		F2	2253	1959	2076	1870	1588	1278	1171	1160	1273	1119	1311
		Dur	93	60	133	68	107	88	133	103	81	95	67
	SD	F1	35	49	32	51	43	84	67	46	47	39	50
		F2	156	120	129	146	74	93	148	145	158	135	93
		Dur	31	18	43	19	30	36	35	36	27	23	20
	Range	F1	153	210	134	255	183	271	290	178	217	167	186
		F2	877	509	606	882	338	409	677	768	815	450	315
		Dur	183	96	232	98	93	158	230	178	118	111	104
	Min	F1	295	356	381	296	553	374	415	378	345	335	494
		F2	1609	1603	1714	1505	1462	1079	932	972	931	932	1180
		Dur	45	32	78	27	58	37	77	47	40	46	29
	Max	F1	448	566	515	551	736	645	705	556	562	502	680
		F2	2486	2112	2320	2387	1800	1488	1609	1740	1746	1382	1495
		Dur	228	128	310	310	237	195	307	225	158	157	133
Female N = 20	Mean	F1	411	496	510	507	805	654	607	541	483	451	689
		F2	2664	2230	2436	2170	1796	1496	1403	1384	1598	1411	1579
		Dur	98	57	136	67	114	100	152	113	75	100	68
	SD	F1	49	64	61	56	68	103	84	69	51	42	64
		F2	186	148	156	138	78	147	164	203	185	165	98
		Dur	31	18	37	21	39	43	71	49	19	32	20
	Range	F1	154	270	261	248	235	531	253	204	182	142	217
		F2	638	416	761	732	274	680	579	623	926	618	327
		Dur	130	74	157	148	170	204	471	235	88	170	102
	Min	F1	335	352	427	352	682	340	486	440	395	395	571
		F2	2301	2007	1879	1980	1661	1176	1118	1102	1245	1040	1414
		Dur	55	24	62	22	59	37	62	34	30	39	35
	Max	F1	489	622	688	600	917	871	739	644	577	537	788
		F2	2939	2423	2640	2712	1935	1856	1697	1725	2171	1658	1741
		Dur	185	98	219	170	229	241	533	269	118	209	137
All N = 40	Mean	F1	388	469	482	482	735	592	568	511	447	425	651
		F2	2459	2095	2256	2020	1692	1387	1287	1272	1435	1265	1445
		Dur	96	58	134	67	111	94	142	108	78	98	67
	SD	F1	42	57	47	53	55	93	76	57	49	40	57
		F2	171	134	143	142	76	120	156	174	172	150	95
		Dur	31	18	40	20	34	39	53	43	23	28	20
	Range	F1	154	240	198	252	209	401	272	191	200	155	202
		F2	758	463	684	807	306	545	628	696	871	534	321
		Dur	157	85	195	123	132	181	351	207	103	141	103

Appendix 5.5. Confusion matrices based on Linear Discriminant Analysis

Predicted vowel type (columns) by intended vowel (rows). Scores are percentages bases on nominally 20 cases per row for native AE speakers, and 120 cases for the PA EFL learners. Automatic classification by LDA with leave-one-out cross-validation when models were trained and tested on American native speaker data (top panel) or with PA EFL learner data. In the bottom panels, the native AE model is used to classify the nonnative tokens (all PA data used for testing). The left part of the table uses two predictors (F1, F2), the right part adds vowel duration as a third predictor. Green cells on main diagonal contain correct predictions (%), red cells contain the most frequent error per row (only when correct prediction ≤ 75%).

		Predicted vowel type, from F1 & F2											Predicted vowel type, from F1, F2 & duration										
		i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u	i	ɪ	e	ɛ	æ	ɑ	ʌ	ɔ	o	ʊ	u
A. American English	i	**100**											**100**										
	ɪ		**85**	15										**100**									
	e		5	**90**	5									15	**85**								
	ɛ				**100**											**100**							
	æ				10	**85**		5									**100**						
	ɑ						55	0	30	15								55	5	30	10		
	ʌ						10	**80**		5	5							10	**80**			10	
	ɔ						45	5	45	5							5	20		70	5		
	o						5	10		60	10	15								5	**75**	5	15
	ʊ				5					15	65	15				5					5	**80**	10
	u											**100**											**100**
B. Palestinian Arabic	i	**88**	4	8									**92**	5	3	1							
	ɪ	4	39	18	39								3	61	3	34							
	e	8	11	74	6	1							6	3	**86**	4	1						
	ɛ	5	26	3	66	0							5	26	2	67							
	æ					**97**		3									**97**	1	2				
	ɑ					0	20	43	11	8	17	1					3	36	23	13	6	18	1
	ʌ					3	9	**82**	4	1	2							13	**84**	1		3	
	ɔ					1	19	18	25	30	3	5					3	12	1	65	15	1	4
	o				1		13	1	23	31	11	20			1			17	2	12	39	10	19
	ʊ				6		1		1	9	62	22				6				2	8	64	20
	u									8	25	67						1		2	7	21	69
C. PA by AE models	i	**82**	6	12									**82**	11	7								
	ɪ	2	65	17	15		2						2	**84**	3	11							
	e	2	14	72	11	1	1						1	6	**89**	2	3						
	ɛ	5	53	5	31		7						4	65	2	28							
	æ					**95**		2	3							2	**93**	4		1			
	ɑ					0	2	29	41	17	8	4					1	40	14	17	10	16	3
	ʌ					8	1	45	44	2						1	1	46	47	3	2	1	
	ɔ					1		25	32	41	1						1	13	1	45	40		
	o						2	17	14	56	5	7			1			21	7	1	57	8	5
	ʊ		2				6	1	1	19	38	34		3						1	19	49	28
	u								1	40	12	46						1			42	9	48