Bangor University

DOCTOR OF PHILOSOPHY

First language attrition and second language acquisition: exploring the role of phonetic aptitude and language use in highly proficient late Arabic-English and English-Arabic bilinguals

Alharbi, Amirah

Award date:
2020

Awarding institution:
Bangor University

PRIFYSGOL BANGOR UNIVERSITY

FIRST LANGUAGE ATTRITION AND SECOND LANGUAGE ACQUISITION: EXPLORING THE ROLE OF PHONETIC APTITUDE AND LANGUAGE USE IN HIGHLY PROFICIENT LATE ARABIC-ENGLISH AND ENGLISH-ARABIC BILINGUALS

Amirah Saud Alharbi

30th April 2019

Supervised by

Dr Anouschka Foltz (1st Supervisor)

Dr Marco Tamburelli (2nd Supervisor)

College of Arts, Humanities and Business, School of Languages, Literatures and Linguistics, Bangor University

A Thesis Submitted in Partial Fulfilment of the Requirements for the degree of Doctorate of Philosophy in Bilingualism (Linguistics) at Bangor University

Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw'r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o'r blaen ar gyfer unrhyw radd, ac nid yw'n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.

I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

Abstract

Few studies have investigated first language (L1) attrition and second language (L2) acquisition of segmental and suprasegmental aspects of speech. The current research includes three separate studies that examined L1 attrition and L2 acquisition of prosody of *wh*-words, vowels formants and the voice onset time (VOT) of voiceless stops in late consecutive Arabic-English (A-E) and English-Arabic (E-A) bilinguals. All three studies explored whether late L2 learners' language use and phonetic aptitude (i.e. talent) affect the prosody of *wh*-words, the formants of shared vowels and the VOTs of voiceless plosives among advanced learners of English and Arabic.

Sixty participants participated in the three studies: 15 E-A bilinguals, 15 A-E bilinguals, 15 monolingual English speakers and 15 monolingual Arabic speakers. Both bilingual groups had been living in the L2 environment for about 20 years and had completed a proficiency test, a phonetic aptitude test and a language background questionnaire.

In study I, the participants read brief dialogues of *wh*-question/answer pairs; the bilinguals read them in Arabic and English. To capture differences in slope steepness and the amount of curvature, the pitch contour of the *wh*-question words was analysed using growth-curve analyses. The monolinguals' results revealed a steep rise to a high target in Arabic *wh*-words, but no such high target in English *wh*-words. The bilinguals' results revealed that A-E and E-A bilinguals approximated the prosodic patterns of the L2. An asymmetrical pattern of L1 attrition was found, showing attrition among the E-A bilinguals, but not among the A-E bilinguals. Additionally, language use, but not phonetic aptitude, modulated how closely bilingual participants approximate native patterns.

In study II, the participants read words containing different vowels that are shared across English and Arabic in isolation and in a carrier phrase. The bilinguals read the vowels in Arabic and English. The first three formants (F1, F2 and F3) were extracted at vowel mid-points using Praat and then normalised. The monolingual groups' results revealed differences in some of the shared vowels in Arabic and English. The bilinguals' results revealed L1 attrition among the A-E bilinguals and the E-A bilinguals. Language use and phonetic aptitude were not found to influence the vowel formants of A-E and E-A bilinguals in L1 attrition and L2 acquisition.

In study III, the participants narrated different cartoons, three in English and two in Arabic, and the bilinguals narrated them in both languages. VOT was measured using Praat from the interval between the plosive release and the onset of voicing. The monolinguals' results showed that the VOTs of voiceless plosives differed in Arabic and English. As in study I, the bilinguals' results showed an asymmetrical pattern of L1 attrition, showing attrition among the E-A bilinguals, but not among the A-E bilinguals. As in study II, language use and phonetic aptitude were not found to influence the VOTs of A-E and E-A bilinguals in L1 attrition and L2 acquisition.

The results suggest that L1 use may prevent L1 phonetic attrition, but only at the suprasegmental level. Some of the segmental results further support the Speech Learning Model (SLM): some areas of pronunciation are more vulnerable to attrition than others and some areas are more likely to achieve native-likeness than others. The present study revealed that some aspects of native-like L2 acquisition can occur alongside L1 attrition, which means that while these aspects are acquired well in the L2, they are susceptible to attrition in L1. However, this is not true for all aspects. Finally, the findings confirm that acquiring a language from birth is not sufficient to ensure L1 native-likeness in the production of bilingual speech.

Acknowledgements

Praise be to Allah (the Almighty) for giving me his blessings and the strength, opportunity and endurance to complete this study.

First and foremost, I would like to thank my supervisor Dr Anouschka Foltz for both her academic and emotional support. I am especially grateful for her guidance in my research and valuable insight into all matters academically and otherwise. She has always supported me, and her direction was key to my success. Thank you so much, Dr Foltz. In addition, I want to extend my thanks to Dr Marco Tamburelli, my second supervisor, for his assistance and support. I also wish to express my gratitude to Professor Ineke Mennen for her supervision during my first year. Her watchful eye provided me with an excellent foundation that allowed me to continue my thesis with confidence. I also appreciate that she kept in touch after moving to another university.

Furthermore, for their complete and generous cooperation and support, I offer my sincere thanks to all the staff at the School of Languages, Literatures and Linguistics at Bangor University, and especially to Dr Eirini Sanoudaki, the Lead of Postgraduate Research. I am also grateful to Dr Sarah Cooper and Dr Marco Tamburelli for all the modules they taught me at the beginning of my PhD, which enriched my knowledge in my area of study, and specifically to Dr Cooper for teaching me how to analyse data using Praat, this study's primary analysis tool.

I would also like to express my gratitude to the Doctoral School at Bangor University for providing me with a range of training (especially the R programme) and development opportunities that enhanced my experience as a postgraduate researcher.

This achievement would not have been possible without the cooperation of my employer, Umm Al Qura University (Makkah, KSA), who granted me the scholarship opportunity to pursue my postgraduate studies in the UK. Credit is also due to the Saudi Cultural Bureau in London for giving me all the support I required.

I wish also to offer my sincerest appreciation and thanks to my colleagues in the Department of English Language at Umm Al Qura University's College of Social Sciences. A special thanks to Umm Al Qura University's Institute of Arabic Language for non-Native Speakers for their contributions to the data collection process. I am grateful to all of the participants who participated in this study.

I could not have finished this study without the full support of my beloved husband (Professor Ramzi Alsaedi) and children (Jehad, Hala and Ayah). Their love, encouragement and continuous prayers have made me stronger every day and have aided me in the completion of this thesis.

Last but not least, my love and gratitude goes to my beloved parents, brothers, sisters and in-laws for their love and continued support throughout the years. I also wish to extend my deepest gratitude to my exceptional friends, Abeer, Wesam, Maram and Amaal. We have supported each other through ups and downs, and your emotional support has eased some of life's most difficult challenges. Finally, my thanks to Mrs. Eirian Siddiqui and her family for looking after my youngest daughter, Ayah, for the past four years. Your assistance has enabled me to work with peace of mind.

Table of Contents

Appendices list

List of Tables

List of Figures

List of Abbreviations

General abbreviations

↑	Indicates the full rise
A-E	Arabic-English
AOA	Age of Arrival
APT	Arabic Proficiency Test
CVC	Consonant Vowel Consonant
E-A	English-Arabic
F0	Fundamental Frequency
F1	First Formant
F2	Second Formant
F3	Third Formant
GCA	Growth Curve Analysis
H*	Simple pitch accent indicating sustained high pitch
IM	Interrogative Marker
L+H*	Complex pitch accent that starts with a low target and then rises to a high target tied to the stressed syllable
L*+H	Complex pitch accent that starts with a low target tied to the stressed syllable and then rises to a high target
L1	First language
L2	Second language
L3	Third language
LOR	Length of Residence
MDH	Markedness Differential Hypothesis
MLAT	Modern Language Aptitude Test
MSA	Modern Standard Arabic
Msec	Millisecond
PLAB	Pimsleur Language Aptitude Battery
SD	Standard Deviation
SLA	Second Language Acquisition
SLM	Speech Learning Model
SSBE	Standard Southern British English
SVO	Subject–Verb–Object
ToBI	Tones and Break Indices
TOEFL	Test of English as a Foreign Language
UK	United Kingdom
US	United States
VOT	Voiced Onset Time
VSO	Verb–Subject–Object
Z3-Z1	Bark-converted F3 minus Bark-converted F1
Z3-Z2	Bark-converted F3 minus Bark-converted F2

Gloss abbreviations

AUX	Auxiliary
FUT	Future
PL	Plural
PST	Past tense
SG	Singular

Chapter One: Introduction

1.1 Introduction

Migration and intercultural contact are increasing worldwide, owing to a range of factors and events, including globalisation and economic pressures and the consequences of conflicts and natural disaster (Liu et al., 2014; Nauck, 2001). Such situations usually require the ability to communicate in more than one language (Nunan, 1999) in order to bridge the gap between different national histories, cultures and social norms. Thus, many researchers have become interested in how migrants and bilinguals generally engage with language: The acquisition of first and new languages has become an important channel of research in the migrant environment (Watson, 1995; Paradis et al., 2011).

Considerable research has been conducted on the first and second languages (L1 and L2) of bilinguals, drawing distinctions between those who learn two languages simultaneously from birth or at a very young age and those who start learning the L2 only in their early teens or later in life. The terminology used in the literature for these different types of bilinguals include (i) simultaneous bilinguals, meaning that both languages are acquired at the same time, and (ii) sequential (consecutive, or successive) bilinguals, meaning that there is a delay between the acquisition of the L1 and the L2. The term 'adult bilingual' is also used to refer to those who have acquired a second language after the age of puberty – often also called second language acquisition (SLA; Lightbown & Spada, 2006).

The participants in this study would be considered 'late consecutive bilinguals' (Hamers & Blanc, 2000; Wei, 2000; de Leeuw, 2008) since they acquired their L2 (either English or Arabic) after the age of puberty, that is, after having fully acquired their L1 (either English or Arabic). The distinction between simultaneous and sequential bilingualism is important in the literature because it is theorised that the level of maturity of the brain at the time of learning has a very significant effect on the outcome (Goetz, 2003). Most of this research has been directed at detecting the influence of the native L1 on the learning of the new L2 in late bilinguals, using terms such as 'borrowing', 'transfer' and 'interference', resting on the so-called critical period hypothesis (Long, 1990; Flege, 1999), which proposes that late language learning is never as effective as early language learning. However, the most recent research argues that it is not appropriate to use absolute terms when theorising about this process: There is great variation among individuals and one can learn a second language at any age (Flege, 2003; Moyer, 2009; Hopp & Schmid, 2013). Some scholars such as Pallier (2007) have noted several critical language acquisition periods that occur throughout one's life, while others prefer the looser term 'sensitive period', which offers more flexibility (Oyama, 1976; Piske et al., 2007).

Research on individuals' varying abilities to acquire another language (Long, 1990; Flege, 1999; Pallier, 2007) has helped develop and refine theories regarding the differences between typical language acquisition settings and backgrounds. Children typically acquire their first language through language input that is received from family members and/or peers, by interacting with others linguistically. Similarly, older children and adults who move to another language community may learn an L2 in a similarly naturalistic setting. The counterpart of such scenarios is the type of structured language learning that older children and adults often undergo in formal contexts (for example, at school or university) or access through different types of devices and media (such as computers and the internet; cf. Cook, 1973; Bialystok, 2001).

More recently, researchers have begun to understand that interference also works backwards, from the L2 to the L1 – in first language attrition (Mennen, 2004). The amount of overlap between an L1 and an L2, which varies according to context (Dorian, 1982), is highly relevant in this respect, highlighting the complex and dynamic nature of second language acquisition (SLA) and language attrition (de Leeuw et al., 2013). Recent studies have started focusing on the language of origin (or first language / L1) in migrant communities rather than on the acquisition of the new language (or second language / L2). Research in this area aims to investigate possible changes in an individual's L1 when they emigrate from their community, as well as to clarify why changes in the L1 occur. L1 attriters are necessarily influenced by their L2, which begins to dominate their thinking when in the L2 environment – the role of L2 acquisition, therefore, is also relevant to the issue of first language attrition. The phenomenon of SLA has received considerably more attention in the scholarly literature than L1 attrition has, simply because there is a widespread monetary and cultural incentive to optimise L2 acquisition in schools and colleges around the world (Nunan, 1999). Most research into SLA is concerned with the L1 only insofar as it inhibits, interferes with or transfers to the acquisition of the L2 (Gass, 1996). In much of the empirical work, the focus is on understanding how people can be helped to acquire an L2 in the most efficient and effective ways possible.

Studies of L1 attrition and L2 acquisition have previously focused on the syntax, lexis and morphology (Sorace, 2004; Tsimpli et al., 2004; Goad, et al. 2003; Köpke, & Genevska-Hanke, 2018; De Clercq, & Housen, 2019, Tsimpli et al., 2009; Leal & Slabakova, 2019). The relationship between the pronunciation of a migrant's L2 and L1 has received less attention. The current investigation therefore focuses on L1 attrition and L2 acquisition in the phonetic domain. Notably, the results of previous investigations with bilingual children (Watson, 1990; Khattab, 2000; Gordeeva, 2006; Fabiano-Smith, & Barlow, 2010; Fabiano-Smith & Goldstein, 2010; Lee & Iverson, 2012; Yang, & Fox, 2017; Al-Amer, 2019) indicate that the phonetic systems of the L1 and the L2 may influence each other (bidirectional influences) in both early and simultaneous bilinguals. Caramazza et al. (1973), Mennen (2004) and Sundara et al. (2006) have proposed that this influence can also be found in late-sequential bilinguals. In the case of the current project, the participants had gained adult competence in their first language (Arabic or English) before they migrated to the UK or Saudi Arabia – i.e. before L2 acquisition. Specifically, because the participants in the current study learned their second language (either English or Arabic) in adulthood after their first language (either English or Arabic), they are considered to be late bilinguals (see Hamers & Blanc, 2000 and Wei, 2000 who differentiate bilinguals in different types). According to de Leeuw et al. (2010), some factors influence first language attrition in the phonetic domain more than others do. Their findings show that the quality of contact with the L1 is often the most important factor in terms of first language attrition. Other factors relating to the way each language is used are also relevant to first language attrition and SLA – for example, it may be that extensive code-switching between the L1 and the L2 leads to greater attrition, as opposed to cases where there are discrete periods during which the L1 and the L2 are spoken separately. Speaking in family, work, study and leisure contexts may influence attrition in different ways. Furthermore, acording to Jilka et al., (2008) and Hopp & Schmid, (2014) phonetic aptitude (i.e. talent) plays a role in both L1 attrition and L2 acquisition.

It is the purpose of the present thesis to narrow the above-mentioned research gap by investigating bilinguals' productions of segmentals and suprasegmentals to explore L1 attrition and L2 acquisition in the phonetic domain. The investigation focuses on the speech of Arabic–English and English–Arabic bilinguals who have lived in the L2 environment for about 20 years. This thesis contains the results from three experiments, which investigate different aspects of SLA and L1 attrition in the phonetic domain. Study I (Chapter Two) concentrates on L2 prosodic acquisition and L1 prosodic attrition in late English–Arabic (E–

A) and Arabic–English (A–E) bilinguals in the production of their *wh*-words. Study II (Chapter Three) examines L2 acquisition and L1 attrition in the segmental productions of the shared vowels of late E–A and A–E bilinguals. Study III (Chapter Four) investigates the voice onset time (VOT) of voiceless stops of late E–A and A–E bilinguals. In addition, all of the three studies explore whether late L2 learners' language use and phonetic aptitude affect how target-like learners are in the prosody of English and Arabic *wh*-words, the production of shared vowels, and the VOT of voiceless plosives. Findings from the current study are organised in accordance with the theoretical implications into two sections – SLA and first language attrition.

1.2 General Overview

1.2.1 L2 Acquisition

Researchers have proposed that pronunciation is one of the most difficult aspects of second language learning (Mennen, 2015; Fraser, 2010). In addition, it has been reported that most of the L2 learners who seem native (for example, in the grammar and other aspects of their L2) still have distinct foreign accents (Flege, 1980, Hinofotis & Baily, 1980; Scovel, 1988; Davis, 1999). For example, Scovel (1969, 1988) argues that many adults with otherwise excellent L2 language skills have difficulties achieving native-like pronunciation in their L2, a phenomenon labelled as 'Joseph Conrad phenomenon'. Joseph Conrad was a Polish-British author, and is one of the greatest English-language novelists. He wrote English extremely well with native-like English grammar and vocabulary, but he spoke English with a strong Polish accent (Scovel, 1969), suggesting that he successfully acquired the English lexicon and morphosyntactic system, but not the English sound system. According to Scovel (1969), many adult learners maintain an accent that reveals their L1. Pronunciation is different from other areas of second language acquisition in that it engages neuro-cognitive-motor skill. One cannot 'learn' pronunciation by simply studying it. Because it involves manipulation of physical tissue, it must be physically practiced, rather like riding a bicycle.

According to Doughty and Long (2003, p.275), "child language acquisition and adult SLA involve different types of processing for language learning". Doughty and Long claim that these differences between adults and children in terms of acquiring a language are signs of maturational constraints. An accent is one of the obvious constraints on an adult's SLA. Indeed, since adults are affected by the speech-processing abilities of the L1, they are considered to be 'disabled' L2 learners. As Doughty and Long (2003, p.284) explain, adult second language learners process speech using a mechanism that is "already attuned to their L1." Long (1990) proposes that if a second language is learned before the age of six, the L2 will be spoken without a foreign accent (maturational constraints), whereas if the L2 is learned after the age of twelve, it will be spoken with a foreign accent. If the L2 is learned between the ages of six and twelve, there will be variability in terms of the presence of a foreign accent. In a similar vein, Patkowski (1990, p.78) points out that there is a 'critical period' or 'sensitive period' for second language learning: Learning a language outside of this time frame is the reason why some L2 learners speak with a foreign accent.

Flege et al. (1999) had participants judge Koreans' pronunciation of English by rating how 'foreign' it sounded. The study then tested their knowledge of English morphosyntax through the use of a 144-item grammaticality judgment test. The study showed a correlational relationship between AOA and degree of a 'foreign' accent exhibited. As these two variables increased, the scores on the grammaticality test showed a steady decline. However, when the confounding variables were accounted for, the relationship between AOA and grammaticality were no longer significant. The accent relationship, though, was maintained. As a result, it can be said that the decrease in morphosyntactic scores was not the result of passing a maturationally defined critical period. Further study suggested that the amount of instruction the Koreans had received in the US (in English) affected generalizable aspects of their English morphosyntax. On the other hand, the amount that the Korean participants actually used English affected the lexical aspects of their English morphosyntax.

In addition, Flege (1999) argues that Patkowski's prediction about the impact of the critical period might be influenced by factors other than pronunciation such as the speaker's grammar and word choice. Flege (1999) presents a review of the literature on L2 acquisition, showing that pronunciation accuracy in the L2 is one of the most difficult skills to achieve, especially if the learner begins L2 acquisition after puberty. A later age of acquisition generally leads to lower pronunciation attainment overall (Flege et al., 2006; Abrahamson & Hyltenstam, 2009). It is often presumed that the age of a person when he or she begins to learn an L2 will correlate with the native-likeness of his or her L2 speech. However, it is difficult to find an objective measure (or indeed a test) of 'nativeness' or 'foreign accentedness' and it is also difficult to separate the age of learning from other factors, such as length of residence or chronological age (Stevens, 2006; Abrahamsson & Hyltenstam, 2009). Munro et al. (1996) have investigated the effects of age on the production of English vowels by native Italian second language learners of English, finding that age is indeed a factor. However, they did also note that there were considerable variations in the ratings that listeners gave to the utterances in the L2 and that there were also variations in the performance of different vowels. These findings suggest that the method is very important when assessing the 'foreignness' of speech production and that there may be particular phonemes that pose difficulties in each language pair that is studied.

Some researchers have questioned whether it is possible to apply the critical period hypothesis more broadly – to the acquisition of L2 phonology in particular (Flege, 1987; Abu-Rabia & Kehat, 2004). For example, Moyer (2009) argues that although age is a factor in L2 acquisition, it is not always the most important factor. Motivation and instruction have also been discussed as predictors of the degree of foreign accent. Motivation is a term generally employed to signify the less-changeable attitudes of students over a long period of time (Moyer, 2009). There are different types of motivation, intrinsic and extrinsic motivation. Intrinsic motivation relates to "motivation to engage in an activity for its own sake" (Pintrich & Schunk, 2002, p. 245), and extrinsic motivation refers to "motivation to engage in an activity as a means to an end" (Pintrich & Schunk, 2002, p. 245). Integrative motivation is the desire to assimilate with the ethnolinguistic group of the language being learnt, whereas instrumental motivation refers to the desire to gain from the learning of the language through employment, social status or through education (Gardner & Lambert, 1959). It has been claimed as well that degree of motivation to speak an L2 is an important determining factor of the degree of foreign accent in the L2 (see, e.g., Suter, 1976; Purcell & Suter, 1980; Bongaerts et al., 1995, 1997; Moyer, 1999).

Lord (2005) observed that suprasegmental instruction in addition to segmental instruction appears to be more effective in achieving a native-like accent than segmental instruction alone (Derwing et al., 1998). Celce-Murcia et al. (2010) as well as Bowen (1972) pointed out that learners of a second language might not effectively be able to transfer aspects of pronunciation learnt through controlled tasks into spontaneous speech. Celce-Murcia et al. (2010) and Hinkel (2006) therefore argue for a communicative element in pronunciation instruction as a necessity to foster fluent and comprehensible L2 speakers.

On the other hand, a study by Bongaerts et al., (1995) found that highly motivated individuals who start learning a second language after the critical or sensitive period can learn to speak an L2 like a native speaker – without a foreign accent. Highly motivated adults – who are supported by good instructors and who avail themselves of opportunities to spend time immersed in the L2 environment – can achieve very high levels of speech performance. This is especially true if training in the suprasegmental aspects of the L2 is offered (t'Hart & Collier, 1975; de Bot & Mailfert, 1982). Although, admittedly, such training is not very common in most instruction-based SLA contexts (Moyer, 2009).

Some previous literature showed that late L2 learners could successfully learn to speak an L2 like a native speaker – without a foreign accent (Bongaerts, 1990; Flege and Fletcher, 1992; Bongaerts et al., 1995; Piller, 2002; Gnevsheva, 2017). Early discussions about pronunciation and accent in SLA was based on the critical period hypothesis, which proposes that early second language learning is a prerequisite for high ultimate attainment levels, and that late learners will generally achieve lower attainment levels, especially in phonology (Bongaerts, 1990). This used to be widely accepted as a typical SLA pattern, but Bongaerts (1990) warns against assuming that this is always the case, suggesting that other factors particular to the character, ability or motivation of the learner, or to the learning context, may also be relevant, and late learners may sometimes achieve near-native accents. In line with this idea, Bongaerts et al. (1995) found native-like English pronunciation in native Dutch speakers.

This recognition of individual differences, and of contextual influences on SLA, led to theories from the domain of sociolinguistics being brought into the debate. According to Piller (2002), findings in parallel areas such as the sociology of gender and ethnicity show that it is neither scientifically nor ethically sound to use very fixed binary categories, such as male and female or native and non-native, to measure phenomena that are socially constructed. Accent is one of the means by which people construct and perform an identity, and it is to be expected, therefore, that a person's accent might consciously or unconsciously change from one context to another, as he or she maintains or projects a particular version of the self for particular purposes. Some people might wish to retain aspects of their L1 phonology in the L2, or alternatively conceal their L1 background for political, emotional or pragmatic reasons, for example.

These debates raise several key issues relating to the methodology that is used to gather data and investigate near-native accent and variation. One implication of the studies mentioned above is that they show how inadequate simple, quantitative measurements of proficiency or attainment level are, because there is no such thing as a baseline or yardstick of native speaker level or accent (Piller, 2002).

This limitation of quantitative and linguistics-based methods prompted researchers to use self-report studies and qualitative analysis to explore contextual factors, as well as the views of second language speakers and their own attitudes towards having an L1-influenced accent in a second language or being able to pass for a native speaker. However, this approach also has some limitations, such as, for example, the inability of speakers to judge how native-like their language is perceived to be, resulting in the likelihood that they will significantly under-estimate or over-estimate their success in passing as a native speaker (Gnevsheva, 2017).

Another, methodological issue is the fact that both speakers and listeners contribute to our understanding of accents and accent perceptions, which means that data should ideally be collected from both these groups. An interesting series of four experiments involving American English native speakers and non-native speakers with L1 Spanish and Chinese was carried out, and the findings showed that even within the space of five minutes, "foreign accent ratings were not stable" (Flege & Fletcher, 1992, p.370). Variability is to be expected, then, in both self-report evidence and accent judgment evidence, and this makes it very difficult to make hard and fast rules about L2 accent levels.

Gnevsheva explores the concept and practice of "passing" which is defined as "an act of being regarded as belonging to a group which one is not a part of" (Gnevsheva, 2017, p.213). Furthermore, Piller (2002, p.179) notes that SLA speakers always describe passing as a "temporary, context- and audience- and medium-specific performance", which highlights the importance of studying the speech of a single speaker across different situations, as well as comparing the speech of multiple native and non-native speakers. Multi-disciplinary and mixed method approaches are therefore best suited to researching this sub-field of SLA.

Previous research has suggested phonetic aptitude has emerged as an important variable in recent work on L2 acquisition. For example, Jilka et al. (2008) found that phonetic talent correlates well with a speaker's performance in spoken tasks. Phonetic aptitude differs from general language aptitude. Phonetic aptitude is discussed in more details in section 1.2.4.

1.2.2 L1 Attrition

Early bilinguals have a better understanding of the morphological and syntactic rules of language compared to late bilinguals (Penfield & Roberts, 1959; Gass & Selinker, 2008) and can adapt to more than one set of socio-cultural norms and assumptions (Chen et al., 2008). In the case of late bilinguals, the research demonstrates that outcomes can be both positive and negative. Studies have documented the difficulty that many late bilinguals have in (i) attaining high levels of competence in the L2 and/or (ii) preserving their competence in the L1 (Krashen, 1981, 1982).

Language attrition occurs when a person loses the ability to speak his or her first (or native) language at native levels. Language attrition is usually caused by lack of contact with other speakers of the first language, learning and frequent use of a second language, or both. A frequently used second language may interfere with producing and comprehending the first language. It is considered likely that attrition of this kind is found to some extent in all speakers of two languages. However, those who use their second language a similar amount to (or more frequently than) their first language will show greater L1 attrition (Köpke & Schimd, 2007).

One definition of first language attrition is ’the non-pathological, non-age related, structural loss of a first language within a consecutive bilingual’ (de Leeuw, 2008, p.10). This explanation concentrates on a language’s functional use rather than on individual language competencies. First language attrition has been defined by researchers such as Köpke (2004) as constituting a loss of a language’s structure, for example, a reduction or change in the individual’s form of the L1. On the other hand, the loss of functional aspects of a language relates to ‘the gradual replacement of one language by another with respect to language use’ (Köpke, 2004, p.4). The individual can experience a decrease in native language ability without experiencing structural loss. This is the situation that occurs for many adults all over the world whenever they may migrate into a new environment and are exposed to a new culture and language after puberty.

Immigrants to other countries where a language other than their L1 is commonly used show L1 attrition. According to Bahrick (1984), L1 attrition may take place in an individual who moves to another country and then stops using their own language completely. Other studies (e.g., Caramazza et al., 1973; Watson, 1990; Khattab, 2000; Gordeeva, 2006; Sundara et al., 2006) seem to indicate that there is a relationship between the phonetic systems of those who learn two languages from an early age, or at the same time.

Scholars have offered many theorisations of language attrition. Using terms such as 'memory' and 'forgetting', Cohen (1975) emphasises the role of cognitive functions. A very common approach is to describe attrition in terms of loss (Lambert & Freed, 1982; Pan & Berko-Gleason, 1986). This terminology implies that the process is unintentional and that there might be some way to locate and revive the elements of the L1 that have disappeared during exposure to an L2 (de Bot & Stoessel, 2000). In fact, the idea of first language loss has been discussed in the literature of children emigrating or being adopted into a new language environment at an early age (Watson, 1990; Hansen, 1999; Khattab, 2000; Ventureyra et al. 2004; Bowers, Mattys, & Gage, 2009). These studies show that even children who have been adopted at an early age and do not remember their birth language retain relevant linguistic knowledge of that birth language (Choi et al., 2017).

More recent research has stressed the importance of the term 'attrition' (Murtagh, 2003; Köpke & Schmid, 2004; de Leeuw, 2009; Mayr et al., 2012), highlighting the different pressures and constraints involved in migrating from one linguistic context to another. The concept of 'forgetting' in relation to language retention has also begun to receive more attention in the field of neurolinguistics, in which a raft of new methodologies has emerged due to recent technological advantages, e.g., functional magnetic resonance imaging (FMRI), that allow a more detailed analysis of attrition to be undertaken (Köpke, 2004).

Early research suggested that the L1 of adults who learned a second language in adulthood is stable and cannot be 'lost' because it is no longer malleable or prone to change (Schmid, 2012). Previous ideas regarding the stability of the L1 considered attrition among adult speakers to be rare, occurring only in specific circumstances. This view has changed, and researchers now acknowledge that influence is bidirectional and that the L1 is far more easily influenced even in late L2 learners than previously thought (Mennen, 2004; Schmid & Köpke, 2014).

Multiple factors are involved in considering L1 attrition. The assumption is often made that simply being exposed to a language is enough to maintain it. This assertion is not, however, supported by evidence (Schimd, 2008). Attrition may be reduced by a positive view of an individual's L1, or the community of people who speak that language, as well as an enthusiasm to maintain the language (Köpke & Schmid, 2004). It is not easy at all, however, to confirm these considerations through research. Adults are more likely to retain their L1 than children are. Therefore, age is certainly a factor in L1 attrition (Dusseldorp & Schmid, 2010; Bylund, 2009; Schmid et el., 2004).

Interestingly, the factors that affect L1 attrition are similar to those that affect L2 acquisition, and parallels can be drawn between the two processes. It must be noted, though, that the impact these factors have on L2 acquisition are much greater than the impact they have on L1 acquisition.

While grammar and pronunciation are retained for a longer period of time, vocabulary that is the first thing those who emigrate as adults forget (Schmid, & Dusseldorp, 2010; Ammerlaan, 1996; Schmid et el., 2004).

Several approaches to language acquisition are particularly consistent with the notion of L1 attrition and the kind of linguistic plasticity that L1 attrition evidences. For example, language acquisition is viewed as emergent and dynamic by usage-based or emergentist approaches. According to Ellis (2016), factors such as frequency, saliency and markedness affect the rate of acquisition.

Holme (2013) and Langacker (1987) explain this plasticity by stating that the brain makes connections when it encounters different phenomena; the more frequently these phenomena are encountered, the stronger the connections are likely to be. Once they have been encountered a sufficient number of times, they gain ‘unit status,’ at which point they will have become a ‘concept’ in the person’s mind. Langacker (2009) suggests that linguistic units are not necessarily stored in this way, but come about as the result of the reoccurring patterns and trends in brain activity. If this is the case, then it stands to reason that the connections associated with a first language may become weakened when a person does not use his or her L1 for a while, or by ‘redirecting’ connections through the conflicting connections of an L2 (Langacker, 1987: 59). Language learning in this context is assumed to be controlled through domain-general learning mechanisms (Langacker, 2009), whose processes depend on the same fundamental ideas of frequency, perceptual saliency, etc. (see Holme, 2013).

Also considered important is the world and context in which the learning finds itself, as well as the incumbent knowledge into which the new information is imbedding itself (MacWhinney, 2012). A child learns its first language, for example, much differently than an adult learns an L2, for many reasons. Their brain, for one, is much more malleable and therefore susceptible to learning during childhood. Children also usually have comprehensive and dependable support from their parents, they connect language to meaning and things they know about the world around them, and they have no previous language onto which they may attach their newly learnt language.

To summarize the current and previous sections, L2 learners often do not achieve native-like L2 pronunciation (Long, 1990; Flege et al., 2006; Abrahamson & Hyltenstam, 2009). In addition, the phonetic system of the L1 may remain flexible in adults and the L2 may influence the L1 (bidirectional influence; cf. Mennen, 2004; Sundara et al., 2006; Stoehr et al., 2017). Thus, learners can stop sounding native-like in their L1 as a result of being exposed to an L2 language environment, either due to exposure to the L2 or due to lack of L1 language use. This will be discussed in more detail in chapters two, three and four.

1.2.3 L2 Speech Models and L1 Attrition

Studies show that L2 learners may encounter issues in any aspect of pronunciation, including consonants (Aoyama, et al., 2004), vowels (Morrison, 2009), and suprasegmentals (Davidson, 2006; Jilka, 2007; Francis, et al., 2008). However, these difficulties do not apply to all learners. For instance, Bongaerts, Mennen and Van der Slik (2000) found that native Dutch learners' pronunciation of English was indistinguishable from an English native control group. These examples of native-like acquisition are rare; however, such examples highlight the likely differences between groups of learners. Studies on the individual differences across learners that affect the degree of foreign accent have found that both the age of the learner at the onset of acquisition as well as the length of residence in their new country affect the acquisition of the L2 sound system (Flege, et al., 2006; MacKay, Flege and Imai, 2006). Other studies highlight different factors, including language use, motivation and other psychosocial factors (Moyer, 2004; Cebrian, 2006). All these studies share the relationship between linguistic and non-linguistic variables.

This section will cover L2 speech models and their relevance to first language (L1) attrition. While these theories do not offer a direct approach towards the issue of L1 attrition, they can however be interpreted accordingly for the purpose of the current study. Within the present project, these models are considered in terms of the way in which they contribute to our understanding of the related phenomena of L2 acquisition and L1 attrition at the level of pronunciation. What is important in this study is the interaction between the two systems in L2 learners, which can take the forms of assimilation (where the properties of the L1 and L2 sounds merge) and dissimilation (where there is a divergence away from the monolingual norm; Flege et al., 2003).

Native Language Magnet (NLM)

According to Tuller et al. (2008), the most important cognitive process that underpins the acquisition of a child's L1 is that of categorisation. In other words, learning a language means learning to perceive differences between sounds and learning to categorise these sounds into meaningful groups. Once a child has formed a strong category around a particular sound, similar sounds are then allocated to that category and, with increasing exposure to the L1, the boundaries between categories grow more secure. This approach has been formally described as the **Native Language Magnet** (NLM) theory (Kuhl, 1991). This model of L1 acquisition explains why infants over the age of around six months begin to discriminate between "prototypic exemplars of phonetic categories in their native language" whereas infants under the age of six months do not (Iverson & Kuhl, 1995, p.123). In effect, the act of listening to the sounds in the L1 environment alters the perceptual systems of an infant so that he or she becomes attuned to the particular phonological system of that language and is able to differentiate between both vowels and consonants, placing them into phonetic categories. What starts out as a general categorisation function gradually becomes language-specific.

The NLM theory was initially developed with a focus on L1 acquisition. But when it is applied to the acquisition of an L2, it postulates that the categorical representations which have been acquired from the L1 will continue to act like a magnet, influencing the way that the learners of an L2 actually perceive stimuli in another language (Kuhl, 1991; 1992). This is bound to happen for two reasons: when learning the L2, (a) there is already an existing and readily available set of categories, which has been developed for the L1; and (b) the individual is older and these early categories have become rather fixed so that it becomes more difficult, if not impossible, to hear the fine differences between some L1 and L2 sounds. In a series of empirical studies, Kuhl and others have established that adult speakers who have different L1 backgrounds perceive the same stimulus from another language differently (Iverson & Kuhl, 1995; Kuhl, 1997; 2004) – they interpret this as evidence of the power of the L1 categories to act as perceptual magnets.

Studies on adult L2 learners have provided a wealth of evidence to support the NLM model since learners consistently experience certain predictable perception and production difficulties when contemplating specific sounds in each pair of languages. Most language teachers quickly become aware of these patterns when they find that students from the same L1 background have the same sort of difficulty with certain phonemes while they manage other phonemes very well. Arabic learners of L2 English have difficulty with the sounds /p/ and /b/, for example, while Japanese learners of L2 English have difficulty with the sounds /l/ and /r/ because these phonemes are categorised differently in the respective languages.

While the NLM model provides useful insights into L2 acquisition, it does not provide any evidence in relation to L1 attrition. Indeed, if anything, it suggests that the L1 categories will remain strong because of their early formation in the developing infant's mind. The NLM model does not illuminate the loss of early L1 categories over time, including over the lifetime of an individual who migrates from one language environment into another. This has led to the development of several other theories that modify the NLM theory and propose new models that explain the effect that learning an L2 has on the L1.

Natural Phonology

There are various explanatory frameworks that highlight the difficulties with L2 speech. In the case of **Natural Phonology**, the difficulty lies in the fact that certain sounds may be considered inherently more difficult than other sounds due to universal constraints. Stampe (1979) notes that the nature of the human articulatory and perceptual systems means that we prefer particular forms. Consequently, if a sound occurs rarely in the world's languages, the human articulatory and perceptual systems do not favour this particular sound. Hence, to learn these forms entails overcoming a perceptual or articulatory difficulty. Natural Phonology argues that adult speakers have adopted a set of natural phonetic constraints and abandoned others. However, each language does have its own set of constraints. There are certain phonological alternations called rules that vary from language to language and need to be learnt. This particular approach does not account for a bidirectional influence between the L1 and the L2.

Markedness Theory

According to **Markedness Theory**, rare sounds are marked, whilst common sounds are unmarked. Markedness Theory would argue that an L2 learner would start acquiring the

L2 through the patterns adopted in their L1. One would expect that it shares its universality as well. This would entail that the attainment of L2 phonology is not only Markedness would predict that the acquisition of sounds that are marked and not present in the L1 would be more difficult to acquire. However, some studies have found language patterns that contradict the relationship suggested by the Markedness Theory (Rice, 2000), and there is also the problem raised by the lack of a generally accepted definition of markedness (Hume, 2004). Besides, the concept of markedness proves problematic. The existence of markedness as a psychological entity is an unfalsifiable hypothesis per se (Rice, 2000). Markedness is a notion that would explain the language faculty that humans have, linking the generativist system to a physical reality. However, there is no evidence for the existence of such faculty. Moreover, there seems to be a lack of agreement on what markedness means (Rice, 2000). For instance, Rice (2000) surveys the literature on markedness and puts together a list of differences between marked and unmarked categories. Marked categories would be less natural, more complex, more specific, less common, unexpected, not basic, less stable, later in language acquisition etc. On the other hand, unmarked categories would be the opposite. Rice (2000) notes that there is cross-linguistic uniformity in the features that can be classified as phonologically unmarked. Conversely, these features are not always the same; thus, there is some local variability. This would entail that the attainment of L2 phonology relies not only on Markedness. Rather, the acquisition of sounds that are marked and not present in the L1 would be more difficult to acquire.

Contrastive Analysis Hypothesis

An alternative approach is to view learners' difficulties as the result of L1 language learning, that is as a consequence of native language transfer. Proponents of this approach hold that the phonological system of the learner's L1 is the starting point for learning the L2. Hence, learners experience difficulties because of the similarities and differences between the L1 and the L2 sound systems. The **Contrastive Analysis Hypothesis** (Lado, 1957) assumes that learners will experience the most difficulties learning the L2 when it differs greatly from the learner's L1. Thus, a systematic comparison between the sound systems of the L1 and the L2 may predict the difficulties learners will face. Hence, according to this approach, phonetic and phonological problems result from two basic reasons: either differences between the mother tongue and the target language, or the interference of the mother tongue. This approach predicts that the rate of attainment of L2 phonology depends on each learner. From this particular point of view, it does not only depend on the L1 and L2, but is also really dependent on the individual. The linguistic analysis offered by this approach may over-predict certain problems that learners face, while under-predicting other problems that learners may encounter, but that are not explained by the transfer between languages. Overall, research has shown that the predictions of the Contrastive Analysis Hypothesis are not always correct (Zobl, 1980).

Speech Learning Model (SLM)

An important and relatively recent theory in this respect, and one which is very relevant to the present study, is the **Speech Learning Model (SLM)**, proposed and later refined by Flege and his colleagues (Flege & Hillebrand, 1984; Flege, 1987; 1995; Flege et al., 2003). The SLM (Flege, 2003; Flege and MacKay, 2004) predicts that L2 sounds that are similar to L1 categories will cause greater difficulties when compared to dissimilar L2 sounds. This model argues that humans tend to assimilate physically similar sounds to the same abstract category. Consequently, a phonetically similar L2 sound will be perceptually assimilated to an L1 category. This phenomenon is known as equivalence classification. Conversely, an L2 sound that is dissimilar to any L1 sound will not be assimilated to native categories and learners will have to create a new category for this L2 sound. In this model, difficulty is assessed by a token by token comparison between the L1 and L2. According to this model, ultimate attainment of the L2 pronunciation is plausible. In fact, Flege (1995: 238) argues that the primary concern of the SLM is the "ultimate attainment of L2 pronunciation". In particular, the SLM suggests that the degree of accuracy in L2 sound perception and production depends partly on the interaction between the L1 and L2 phonetic systems of a learner. This approach also notes that age plays a key role in ultimate attainment of L2 speech. For instance, early learners, that is those who start L2 speech learning before they reach the age of 12, are more likely to outperform those who start learning at a later stage. This particularly occurs in terms of perception and production of L2 vowels (Piske et al. 2002) and consonants (MacKay et al. 2001). Similarly, early learners tend to be judged as having less strong foreign accents (Flege et al. 1995). This particular approach allows for a bidirectional influence between the L1 and the L2, as Flege's (1995) model assumes that the speech sounds of any two languages are represented in a shared phonological space. The idea that they are in the same space implies that there may be a bidirectional influence between the sounds of both languages. This, in turn, predicts that there may be articulatory changes in the L1, where the learner is copying the articulatory characteristics of the L2. In other words, sharing a space may result in the assimilation of the articulation of individual sounds from the L1 to the L2.

The SLM aims to account for ultimate attainment, "so work carried out within its framework focuses on bilinguals who have spoken their L2 for many years, not beginners" (Flege, 1995: p. 238). The aim of the SLM is to explain how speech sounds in the L2 are acquired. It also contributes to our understanding of how this process in turn affects the L1. Many studies of bilingual migrants have explored what happens when speakers encounter sounds in the L2 that are different from those of their L1. There are studies which focus on the period of initial contact (Flege et al., 2003), and others which focus on the ultimate attainment of learners (Bongaerts et al., 1997) and still others which attempt to trace the development that occurs in phonological perception and production over long periods of time (de Bot & Clyne, 1997; Cook, 2003).

Similar to the Contrastive Analysis Hypothesis, the SLM highlights the importance of the degree of difference between sounds. There are two routes through which first and second language speech sounds can interact. The first, assimilation, occurs when no new sound category in the second language can be established. However, the SLM proposes that the most *similar* sounds are the most difficult to learn for L2 learners. In more detail, learners often fail to distinguish small differences and, in this case, a new L2 sound is simply assimilated into the existing L1 category. A sound that is similar (i.e. an L2 sound which is similar to the L1 sound) is likely to be categorized into the same category as that of the L1. Therefore, when an L2 sound is considered similar enough to an L1 sound, it will be merged (or assimilated) with it. Alternatively, dissimilation (or polarization) happens when a new category is perceived in a second language speech sound. Thus, if the learner perceives a difference between the L1 and L2 sounds, then a new category is formed; this process is called 'category formation' (Flege, 1987; Flege, 1995; Flege et al., 2003; de Leeuw, 2009). Bilinguals seek to find differences between corresponding L1 and L2 sounds, just like monolinguals seek to establish a contrast between similar sounds in their L1. This process causes category dissimilation (Flege et al., 2003). In summary, according to the SLM, in attrition of a first language, and acquisition of a second language, one of two things may happen when similar sounds are encountered (depending on whether new categories are formed or not): assimilation or dissimilation.

The SLM helps us to explore these two contrasting strategies, both of which are likely to occur in the process of acquiring an L2. It can help to predict which learners are likely to use which of these strategies and which sounds are most likely to undergo assimilation vs. category formation in particular language pairs. In turn, this can inform the design of language teaching programmes for different learner groups. From empirical studies on this point, it is clear that "all else being equal, early bilinguals will be more likely to establish new phonetic categories for L2 speech than late bilinguals will be" (Flege et al., 2003, p.469). How far these choices then become part of a shared phonological system across groups of L2 speakers who have the same L1 background and therefore very similar assimilated or polarised sound categories is a question that remains to be investigated. It could be, for example, that frequent contact with other speakers who share the same bilingual background and history promotes the formation and retention of hybrid categories as opposed to distinct L1 or L2 categories.

Importantly for first language attrition research, the SLM predicts that "the more a bilingual approximates the phonetic norm for an L2 speech sound, the more her production of the corresponding L1 speech sound will tend to diverge from the L1 phonetic norm" (Flege et al., 2003: p. 470), while this effect is stronger in late bilinguals than in early ones. This suggests that the more accurate the approximation of the L2 speech sounds, the further the equivalent L1 production will stray from the standard sound (Flege et al., 2003). This effect is, however, magnified in late bilinguals.

According to the SLM, the main predictor – including age – of a person being able to establish a new phonic category for a sound in their L2 is their own ability to perceive a sound to be adequately different to the equivalent L1 sound. However, age does have some impact on how the L1 and L2 phonetic sub-systems interact. "That is, as L1 vowels and consonants develop, they will perceptually assimilate neighbouring L2 vowels and consonants more strongly" (Flege et al., 2003: p. 469).

The SLM predicts that L2 learners' phonetic categories may differ from those of L1 monolinguals even if the bilinguals or learners have been learning the L2 for a long period of time. The perception and production of sounds will be affected by such differences in the stored categories in the mind. Primarily, the SLM refers to phonetic categories and not phonemic ones. For example, Flege holds the English [t] and [t^h] as being categorised differently and put forward that native speakers have different mental representations for them since they differ phonetically (e.g. in VOT). According to the SLM, English native speakers do not have one category for the English [t] and [t^h] just because [t] and [t^h] are allophones of the phoneme /t/, i.e. allophones of the same phoneme.

Apart from situations where sounds in both languages are the 'same' (at the allophonic and phonetic level), the SLM assumes that some distinctive types of situations can arise. This leads to different learning routes as L2 segments are learned; those for 'new' and various types of 'similar' sounds (Flege, 1987c; 1995). Flege clarifies this theory along the lines of certain L2 sounds being perceived as either 'new' or 'similar' to the learners and to the learners' native language (L1). Following this, the IPA symbol criterion was suggested as a rough guide to understanding this difference (Flege, 1992). Evidence of this is found in the English [p] since no phone in Arabic would be labelled with the IPA symbol [p] and it is therefore new to the Arabic learner. However, both languages have a phone that would normally be written as [b] in the IPA, making [b] similar. For Flege, however, the English and Arabic [b] would not equate to the 'same' sound because of the differences reflected in the VOT (with Arabic in general being pre-voiced and English having a short-lag positive VOT). According to Flege, similarity is defined by the allophonic properties of the sounds, and not by the broad phonemic correspondence (cf. the Contrastive Analysis Hypothesis (CAH); e.g. Lado, 1957).

Another of Flege's claims is that a person's ability to learn new sounds remains active and develops throughout their life (Flege, 1995), contrasting with the critical period hypothesis that states that learning new sounds after the age of puberty would be extremely difficult if not impossible (Scovel, 1988). Although Flege does agree with this to a certain degree – that the learnability of L2 sounds decreases as the age of learning increases – he disagrees with the idea of the complete loss of the ability to acquire new sounds after the critical period. Rather, Flege proposes the idea that while in the early stages of learning, but regardless of age, learners should filter L2 sounds through the L1 sound system. Flege (2003) adds that the filtering in the later stages can be halted if extensive exposure to the L2 sounds is maintained, allowing the learner to establish new L2 categories.

Flege (1995, p.239) proposed the following four main postulates:

1. The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span and can be applied to L2 learning.
2. Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.
3. Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as the realization of each category.
4. Bilinguals strive to maintain contrast between L1 and L2 phonetic categories that exist in a common phonological space.

Ontogeny Model

Flege's (1993) account is a theory that combines universal factors and cross-linguistic influence. There is a further model that combines these two: Major's **Ontogeny Model** (Major, 1987). This particular approach on the acquisition of L2 phonology addresses the competing influence of developmental factors and cross-linguistic influence. Major (2001) developed this model as the Ontogeny Phylogeny model. Within this approach, the premise established by Major (2001) is that cross-linguistic factors decrease over time. At the same time, the influence of developmental factors, such as the effect of the universal difficulty of particular sounds first increases and over time begins to decrease again. To explain this particular pattern, Major (2001) proposes that there is an interaction between the similarity of the L2 to the L1 and also markedness factors. That is, learners of two languages with more similar sounds are more strongly affected by cross-linguistic influence. Increasing levels of markedness relate to a higher influence of developmental factors instead. This model makes rather strong predictions regarding L2 pronunciation. However, there are no studies that have empirically explored this particular approach, unlike Flege's (2003) model, which has been put to the test in numerous studies.

Perceptual Assimilation Model (PAM) extension (PAM-L2)

Alternatively, the **Perceptual Assimilation Model** (PAM) **extension (**PAM-L2) (Best, 1995; Best & Tyler, 2007) bases difficulty on the assimilability of non-native contrasts to native categories. SLM and PAM/PAM-L2 "base their predictions of the relative difficulty or ease of production and perception of non-native speech on comparisons of L1 and the to-be-learned *segments"* (Mennen, 2014, p.6). PAM-L2 as well as Flege's (1993) theory share a main premise. Both models assume that learners may perceive sounds as functionally equivalent when these L2 and L1 sounds are phonologically or phonetically similar. The central assumption of PAM is that speakers assimilate non-native sounds to the native sounds they perceive to be more similar. As such, perceptual similarity is defined with regards to dynamic articulatory information; i.e., how articulatory gestures shape a sound. Hence, the accuracy to discriminate non-native sounds is strongly dependent on how these sounds are assimilated to the mother tongue of the learner. That is, learners are more likely to distinguish L2 contrasts depending on the articulatory settings of the L1 and the L2.

Second Language Linguistic Perception Model (L2LP)

Drawing on elements of this theory, Escudero (2005) argued for the **Second Language Linguistic Perception Model** (L2LP). In order to capture the whole developmental process of L2 speech perception, this model suggests that L2 learners initially perceive target language sounds in the same way as they perceive their native language sounds. However, studies employing this approach have shown that it is possible that learners of an L2 alter their articulatory perception of the L1, resulting in sounds between the speech sounds of the two languages (Escudero & Boersma, 2002; Escudero, 2005, 2009). The model proposes the Full Copying Hypothesis that states that learners create a duplicate of their L1 system when they first encounter the L2. Hence, L2 perceptual development does not affect the L1 system, meaning that transfer can only occur at the onset of L2 learning (Escudero & Boersma, 2004). The model also posits that sound perception is shaped by the acoustic properties of the individual accent of the learner's L1, such that individuals with the same L1 may have differences due to their accents. This means that L2 development is contingent on the individual ability of the learner to overcome their particular problems.

For instance, a particular learner may initially perceive the two sounds of the L2 contrast in terms of a single native category, assimilating the two sounds as one. In this case, the learner faces a different set of learning tasks than a learner who assigns the same L2 sounds to two separate L1 categories because of their accent. If the learner assimilates the two sounds as one, the learner has to either create a new L2 category or split the single category handling both elements that contrast. On the other hand, if a learner assigns two separate categories, the learner does *not* need to create a new category; the learner only needs to *reuse* their existing L1 categories and shift their L1 perceptual boundary so it matches with the L2. The second scenario presents an easier task, and examples of L2 Category Boundary Shifts have been found in the literature (Escudero & Boersma, 2002; Escudero, 2009). In other words, the L2LP model aims to account for the entirety of L2 perceptual development. In regards to ultimate attainment, Escudero and Boersma (2002) note that native-like ultimate attainment may be possible in certain circumstances. However, if a learner's L1 contrasts are left intact as they acquire an L2 without this contrast, this may hamper the attainment of a native-like command of the L2.

L2 Intonation Learning theory (LILt)

The approaches and theories mentioned so far have mostly focused on segments. However, an area of pronunciation that is particularly problematic for L2 learners is intonation. L2 learners almost always speak with intonation patterns that differ from those of native speakers, even with years of practice and exposure (Jilka, 2000; Mennen, 2004; Munro, 1995; Trofimovich and Baker, 2006; Willems, 1982). This alone can lead to the perception of a foreign accent (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Jilka, 2000a; Mennen, 2004; Magen, 1998; Munro, 1995; Munro & Derwing, 1995; Trofimovich & Baker, 2006; Willems, 1982). Some believe that intonation is more susceptible to cross-language influences than other aspects of pronunciation (Mackay, 2000). As a result, it is often found that even in highly proficient L2 speakers, their L1 will affect their L2 intonation (see Mennen, 2004, 2007 for an overview). Segmental acquisition makes up most of the research in the study of L2 speech production and perception, suggesting that there is quite a good understanding of the segmental aspects of language differentiation. Due to this, Flege's SLM (1995) and Best's PAM/PAM-L2 (Best 1995, Best & Tyler 2007), amongst other theories of L2 speech production, compare the ease-of-acquisition of L2 to L1 and the to-be-learned *segments*. To date, the **L2 Intonation Learning theory** (LILt) is the only model that deals with and makes predictions regarding the relative difficulty of generating, producing and distinguishing intonation of the non-native variety. There have, however, been some attempts made to extend the PAM-L2 so that it includes the perception of lexical tones (So & Best, 2008, 2010, 2011).

LILt "agrees with SLM and PAM-L2 that many difficulties may be perceptually motivated, it posits that explicit reference needs to be made to the semantic dimension of intonation when determining perceptual similarity. Finally, as with segmental models, the

LILt does *not* rule out other explanations of deviations in production, such as an inability to articulate certain differences between L1 and L2 intonation or store them in acoustic memory" (Mennen, 2015, p. 197). The theory requires that the position and context of certain contrasts be tested and controlled for, as they are also important in regards to intonation. LILT holds that, as learners increase their experience of the L2, their intonation parameters will more closely approximate the norms of the L2. Learners use their L1 to produce intonation in the L2 in the absence of L2 experience. This use of L1 intonation in the L2 language is seen more often at the early learning stages (e.g. McGory, 1997; Mennen, 2004; Jun & Oh, 2000; Ueyama & Jun, 1998). LILt assumes a bidirectional influence, so it could account for both L2 acquisition affects and L1 attrition affects.

Out of the models reviewed in this section, the SLM approach and LILt are the most useful for the present study. The SLM and PAM-L2 approaches are most useful for segments because many previous studies have used the framework, so that it makes comparison with previous empirical findings easier. Moreover, it accounts for a bidirectional influence. In addition, LILt is most useful for prosody since it is the only model to date that makes predictions of and deals with the difficulty of perceiving and producing non-native intonation. Also, LILt assumes a bidirectional influence, so it could account for both L2 acquisition affects and L1 attrition affects.

There are still some methodological problems, which make SLM difficult to apply in practice, however. There is a lack of consensus in the literature on how to determine what constitutes similar and dissimilar sounds and this is a major weakness in the theory (Bohn, 2002). The SLM is also a somewhat static model, which does not adequately account for the high level of variation that occurs between speakers who have the same L1 and L2 pair or even in the same speakers in different contexts and during different life stages. Some L2 learners are able to overcome initial sound perception problems and shift from an assimilation strategy to a polarisation strategy either through intensive tuition, long residence in an L2 environment or due to other, unknown factors. Others, on the other hand, experience a fairly constant retention or steady attrition of the L1 sounds. These dynamic and variable features of speech perception and production are not easily accounted for using the SLM approach (de Leeuw, 2009).

In light of the partial explanations offered by the multiple theories reviewed above, the present study endorses the suggestion made by de Leeuw (2009, p.38) that "a multicomponential theory of first language attrition" is what is required. No single model explains all the data that researchers have found in their fieldwork. Insights from the models mentioned here provide a valuable theoretical background but there is still some work to be done in accounting for all the phenomena that are encountered in studies of L1 acquisition, L2 acquisition and L1 attrition, as well as the complex relationships which can exist between these three related processes.

Finally, there is a further problem that has thus far been only partially addressed by the range of L2 speech models available. This problem can be summed up as a failure to take account of the settings in which L2 learning and L1 attrition occur. This is a vital dimension which must be included if researchers are to be able to determine even the simplest kinds of cause-and-effect relationships. It is entirely possible that other factors beyond speakers' cognitive processing could be responsible for the multiple variations that have been recorded in the scholarly literature. If no attempt is made to measure these factors, then doubt must be cast on any conclusions that are reached. Research into models for L2 acquisition and L1 attrition is continuing at the present time and there is a growing awareness of the need to include these factors as well (Major, 1992; Mennen, 2004; de Leeuw, 2009). The present study hopes to contribute to this theory-building work as well as to conduct empirical research into the L2 speech and L1 attrition of bilinguals.

1.2.4 Predictor variables

L2 speech models were explored in the previous section. In this section, the general focus is on predictor variables that may influence the process of phonetic L2 acquisition and L1 attrition. There are several factors that may affect both L2 acquisition and L1 attrition. For example, Köpke and Schmid listed the following predictor variables which have been found to have an impact on the degree of L1 attrition: "age at onset of L2 acquisition, age at onset of L1 attrition, time since onset of attrition, level of education, attitudes, frequency, amount and settings of use of the attriting language" (2004, p.15).

It has long been thought that length of residence (LOR) has a strong influence on degree of foreign accent. In addition, the longer an individual resides in a new location where their L1 is not the native language, the more likely s/he will experience L1 attrition (Lambert & Freed, 1982). For example, Schmid (2004) showed that, when compared with monolingual speakers in Turkey, Turkish immigrants in Australia who had lived in their new country for a long time had certain difficulties in adequately processing and/or producing relative clauses in Turkish. Rather than using relative clauses, L1 attriters often adopted an alternative strategy in which two dependent main clauses are combined. Such an outcome is a good example of syntactic and grammatical attrition. This precedent illustrates the consequences of a lack of adequate stimulation on the L1 (cf. e.g. Paradis, 2007).

The frequency and type of L1 used in an L2 context has been identified as another relevant factor in L1 attrition and L2 acquisition, particularly with respect to LOR. According to Schmid (2007), it is possible to trace different patterns of L1 attrition in speakers who use the L1 only at home within the family and those who have wide connections with other L1 speakers, such as an expatriate community that celebrates its L1 heritage. It is fair to conclude that the contexts in which the L1 is used may substantially influence the degree of L1 attrition observed in individual speakers, which may interact with the regularity of L1 use (or lack thereof).

Schmid (2007) affirmed that different types of experimental tasks and the participants' levels of education (EDU) may elicit different types of effects associated with LOR. For example, in the case of German-Dutch and German-English bilinguals, verbal fluency and lexical retrieval were shown to have a significant impact on the degree to which speakers were able to successfully retrieve lexical items from a semantic field (such as animals and fruits/vegetables). Story-retelling tasks and situations where free speech is elicited seem to have little or no effect on L1 attrition.

The findings on the frequency and type of L1 used in the L2 context therefore reflect a range of factors. Such circumstances include not only individual variability in performance, but also the degree to which EDU and different types of training or linguistic experience in formal contexts allow speakers to retain L1 fluency in different linguistic tasks. The degree to which connections are shared with other L1 speakers in a community interacts with those types of variables, which highlights the dynamic nature of how L1 attrition is related to both the frequency and type of L1 and L2 usage and LOR (Schmid, 2007).

In situations where negative attitudes towards the L1 in a majority L2 culture pervade a language community, that type of influence can accelerate L1 attrition in individual speakers due to their desire to blend in with the dominant culture. However, this can also reinforce L1 retention in certain individuals to the extent that they wish to resist the dominant culture (Pavlenko, 2003, 2005, 2008). Such social and cultural factors are variable among individual speakers, even within the same family or the same community.

More specifically, LOR may have a range of effects on L1 attrition, including various types of L1 loss, whether this loss is structural, semantic, pragmatic or interactional. Overall, in terms of acquisition and attrition, the situation seems more complex in this respect than has been suggested by some early studies (such as Cohen, 1975; Chomsky, 1969; Grosjean, 1982). A multitude of factors affect the outcome of any given L1 attrition situation, including social and linguistic factors, as well as issues specific to individual speakers and the distinctive networks they share with other L1 and L2 speakers.

Primarily to facilitate a discussion on this point, de Leeuw (2009) warned against making narrow assumptions about the acquisition of the phonetic system of the L2 being a direct cause of the attrition of parts of the phonetic system of the L1. This is because there are many other factors, apart from purely phonological considerations, which are bound to affect people in many different ways. Indeed, according to de Leeuw (2009), a great many variables can predict the attrition of the L1: the frequency and quality of language contact, including aptitude, language attitudes and sex, as well as the phonological properties of each language. This means that it is possible to trace patterns and make connections between L1 and L2 phonology, and it may even be possible to identify some segments that are clearly attributable to the interference of the L1 on the L2 or the attrition of the L1 due to the intrusion of the L2 – there are, however, bound to be other segments which are not affected in this way. There is also likely to be great variation between individual speakers due to the diversity of variables in the environment and within each individual.

Previous studies have examined the influences on L2 acquisition according to the length of residence in an L2 country or the amount of L1 and L2 use (Trofimovich & Baker, 2006). The impact of language use and length of residence on L2 acquisition and L1 attrition have also been examined in relation to learners' knowledge of segmentals and suprasegmentals (de Leeuw, 2009; Flege & Fletcher, 1992; Guion et al., 2000; Leather, 1987; Mayr et al., 2012; Moyer, 1999; Purcell & Suter, 1980; Shen, 1990; Thompson, 1991). Despite a substantial body of research, very few studies have looked specifically at everyday L1 and L2 language use or the role of the quantity and quality of language use when maintaining native-like L1 speech patterns after moving to an L2 environment (de Leeuw et al., 2009; Mayr et al., 2012).

Studies have shown that age of arrival in migrant populations affects language choice, with early L2 learners (with ages of arrival younger than 10) primarily communicating in the L2 and late L2 learners using their L1 more frequently when talking to members of their home/native community, as well as in the target-language environment (cf. Jia, Aaronson, & Wu, 2002; Piske, MacKay, & Flege, 2001). Age of arrival also has an influence on L2 acquisition and L1 attrition with more native-like L2 productions as well as more L1 attrition in earlier arrivals in comparison to later ones (Flege et al., 1995; Flege, MacKay, & Meador, 1999; Lee, 2011).

Since it can be more difficult to objectively measure the amount of language use in an average day or week than a learner's age of arrival in migrant populations, subjective measures (such as learners' statements and reports) are often used to determine how often bilinguals use their L1 and L2 on a daily basis. This method holds a risk of inaccuracies in participants' reporting, which may have been why Flege (2009) reported that the quantity of L1/L2 use accounted for less than 10% of the difference in foreign accent rating data in a meta-analysis of data from 240 L1 Korean and 240 L1 Italian L2 learners of English across several tasks and studies. Alternatively, it may not be the quantity but the quality of L1 and L2 use that has a greater effect on L2 acquisition and L1 attrition. For example, neurobiological research on language learning and stimulus appraisal has implied that emotional involvement might be necessary for effective L2 learning (Schumann, 1998). Furthermore, engagement in conversations with native German speakers in German aided L2 learners of German in their acquisition (Hopp & Schmidt, 2010). This idea is further supported by Suter's (1976) and Purcell and Suter's (1980) early research into the influence of language use on the degree of a foreign accent in L2 learners. In their studies' questionnaires, the learners of English estimated how much they used English at work, home, and in school with native speakers of English as well as how much time they spent with English native speakers. Cohabitation with English native speakers (i.e., more L2 use in the home) was identified as being the third most predictive variable. The amount of L2 use at work or school were less predictive due to their correlation with L1 background and length of residence. However, Flege & Fletcher (1992) found no significant correlation between the percentage of everyday L2 English use and the degree of an L2 foreign accent in their native Spanish-speaking participants.

Thompson (1991) asked L1 Russian immigrants to estimate how often they used English at work, at home, and with friends. A simple and significant correlation between the amount of English use and the degree of an L2 foreign accent was found. However, an analysis of multiple regression did not confirm the amount of English use as being a significant predictor (since it was confounded with age of learning). Since some immigrants with high L1 proficiency had an L2 foreign accent, the hypothesis for the study was that the L1 had a possible effect on L2 pronunciation. Based on this, the researcher argued that "a difference must be noted between subjects who have maintained their mother tongue and those who have lost it when it comes to estimating accent retention in the second language" (Thompson, 1991, p. 200).

Researchers into L1 attrition such as Schmid (2007) and De Leeuw et al. (2009) examined the impact of language use. De Leeuw et al. (2010) sought to determine whether native German speakers residing in the Netherlands or (Anglophone) Canada are perceived as having a foreign accent in their L1 German. Nineteen German monolingual listeners assessed the foreign accents of five German monolinguals as well as 34 L1 German speakers residing in Canada and 23 L1 German speakers living in the Netherlands (who had migrated at an average age of 27 and who had been living there for averagely 37 years). The results showed that nine of the immigrants to Canada and five of the immigrants to the Netherlands were perceived as being non-native German speakers. De Leeuw et al. (2010) concluded that the quality and quantity of language contact with German native speakers had a greater influence on predicting a global foreign accent in native speech than the length of residence or age of arrival. In particular, communicative settings where code-switching was unlikely to occur predicted foreign accent ratings of the L1 speech, but the amount of L1 use in communicative settings where code-switching was likely to occur did not. Thus, speaking the L1 in settings where code-switching is unlikely to occur seems to be conducive to maintaining a high level of native language pronunciation within consecutive bilinguals in a migrant context.

Generally, it appears that certain kinds of language use, for example in the home or in situations where code-switching is unlikely, have an influence on L1 attrition. It has to be noted that most of these studies concern a global foreign accent, so it is not clear which aspects of pronunciation may be affected by language use.

Phonetic aptitude (or phonetic talent) has not received as much focus in previous studies as variables such as LOR and language use, which were described above.

Nevertheless, phonetic aptitude has emerged as an important variable in recent work on L2 acquisition. For example, Jilka et al. (2008) found that phonetic talent correlates well with a speaker's performance in spoken tasks. Phonetic aptitude differs from general language aptitude. General aptitude was defined by Carroll (1981, p.1) as the "capability of learning a task", which depends on "some combination of more or less enduring characteristics of the learner". In addition, general language aptitude is a talent for all aspects of language, such as vocabulary and syntax. Moreover, Nardo and Reiterer (2009) defined *talent* as "i) a characteristic feature, aptitude, or disposition; ii) the natural endowments of a person; iii) a special, often creative or artistic aptitude; iv) general intelligence or mental power (ability). *Aptitude* can be characterized as: i) an inclination or tendency; ii) a natural ability (talent); iii) a capacity for learning; iv) a general suitability (aptness)" (p. 213).

Phonetic aptitude, on the other hand, refers to a person's talent at pronouncing the foreign/second language. Jilka et al. (2011, p.171) define phonetic talent as "an innate, somewhat mysterious ability that a person either has or does not have". Notwithstanding, the term "aptitude" is open to interpretation. It is therefore very important in any study focusing on this factor to provide detailed and adequate definitions of the properties of, and the issues related to, phonetic aptitude, as well as to devise reliable measurements for phonetic talent.

The approach taken by Jilka et al. (2008) informs the present study. It consists of a range of tests, some of which are specific to the languages investigated and some of which are divorced from that context, such as more abstract tasks involving the acquisition and reproduction of items in a language that the participants do not know. The use of several different types of instruments increases the reliability of the results and may help to delimit aptitude from other factors, such as the amount of instruction received or the extent of linguistic experience. The work by Mennen (2004) highlighted the fact that some participants in L2 acquisition tests can be seen as "outliers" from the main trend. This issue raises the question: What makes these individuals particularly good at learning the L2? Some participants in Mennen et al.'s (2004) study exhibited no cross-linguistic interference in formal reading in the L2. This finding suggests that participants who do well in the L2 yet show no effects on the L1 may be especially phonetically talented and that this skill explains their success. Language use and phonetic aptitude will be explained in greater detail in chapters two, three and four.

1.3 Statement of the Problem

In the past, the focus in much of the empirical work has been on understanding how an L2 can be acquired in the most efficient and effective way. This body of literature is included briefly in this study, but the main focus of the present study is on L1 attrition and the question of the relationship between the L1 and the L2. Research into L1 attrition has, during its short history, been focused on some of the linguistic domains – for example, syntax. Migrants' L1 pronunciation and the relationship between late L1 attrition and L2 acquisition in the phonetic domain have received very little attention. In order to narrow this gap, the present study seeks to make a direct comparison between late L1 attrition and L2 acquisition in the phonetic domain and explores how the amount of L2 language use and phonetic aptitude influence both L1 attrition and L2 acquisition in segmentals (shared *vowels* and *VOT* of the voiceless plosives) and suprasegmentals (intonation of *wh-* words). In doing so, the current project focuses on the productions of highly fluent, late consecutive A–E and E–A bilinguals. In addition, the current project attempts to test whether conclusions drawn from previous studies will hold true across two typologically different language groups such as Arabic and English.

A better understanding of the factors involved in L1 attrition is important as it will illuminate the processes of language loss and learning within the brain and could even contribute to the development of pedagogies that can succeed in optimising L2 acquisition without concomitant L1 attrition. In addition, it will contribute to the debate regarding the biological and social influences that affect bilingual adults, as well as having practical implications for such bilinguals as they interact with L1 and L2 speaker networks throughout their lives. It may be possible, for example, to predict which phonological difficulties might arise in particular groups of learners and then to design appropriate instruction methods to remedy these difficulties in the classroom (Eckman, 1981). With these aims in mind, this project focuses on the domain of phonetics and the relationship between L1 attrition and L2 acquisition in highly fluent and late consecutive A–E and E–A bilinguals.

1.4 Research Questions

This project is designed to explore L1 attrition and L2 acquisition in the production of *wh*-words, formants of shared *vowels* and *VOT* values in *voiceless stops* in late bilinguals in addition to exploring whether late L2 learners' language use and phonetic aptitude affect how target-like they are in the production of segmentals (*vowels* and *VOT* of voiceless plosives) and suprasegmentals (prosody of *wh*-words) of both Arabic and English. Language use seems to be especially related to L1 attrition. Indeed, previous researchers (Mayr et al., 2012; Stoehr et al., 2017) have suggested that prolonged use of, and exposure to, the L2 is needed for speakers' L1 pronunciation to attrite. Phonetic aptitude was also selected as a factor to investigate because its relationship to L2 acquisition and L1 attrition is unclear and has to date received relatively little attention.

In detail, the present work aims to address the following research questions:

- Do highly fluent late consecutive Arabic-English and English-Arabic bilinguals who moved to an L2 country show native-like L2 productions of segmentals (*vowels* and *VOT* of voiceless plosives) and suprasegmentals (prosody of *wh*-words) and/or do they show attrition in their L1 productions of segmentals (*vowels* and *VOT* of voiceless plosives) and suprasegmentals (prosody of *wh*-words)?
- Does language use and phonetic aptitude play a role in L1 attrition and L2 acquisition of segmentals (*vowels* and *VOT* of voiceless plosives) and suprasegmentals (prosody of *wh*-words)?

1.5 Organisation of the Study

The main chapters of the current thesis are not written as traditional thesis chapters, but are instead written to resemble journal papers. Chapters Two, Three and Four each present a separate study that is written up in the style of a journal paper, including a separate abstract, introduction, methods section, results section and discussion. Writing up the thesis in this way was done to facilitate future publications of the results to peer-reviewed international journals.

The current chapter is a brief, general overview of major themes in L1 attrition and L2 acquisition research. Moreover, the statement of the problem and the research questions are also reported in the current chapter. Chapter Two will discuss L2 prosodic acquisition and L1 prosodic attrition in late E–A and A–E bilinguals in the production of their *wh*-words. Chapter Three investigates L2 acquisition and L1 attrition in the shared vowel productions of late E–A and A–E bilinguals. Chapter Four will discuss the production of VOT by monolingual native speakers of both languages, as well as the relationship between L1 attrition and L2 acquisition in highly fluent and late consecutive A–E and E–A bilinguals. In addition, all three chapters (Two, Three and Four) will investigate whether late L2 learners' phonetic aptitude and language use affect how target-like learners perform in terms of their *wh*-words, shared vowels and voiceless plosives of both Arabic and English. Finally, Chapter Five summarises the results from the three previous chapters and presents a general discussion of the results, including their limitations and the implications of the findings. Recommendations and suggestions for future research are also presented in Chapter Five.

Chapter Two (Study I)

L2 acquisition and L1 attrition of prosody in highly-proficient late bilinguals: Exploring the role of phonetic aptitude and language use

Abstract

Relatively few studies have examined the second-language (L2) acquisition and first-language (L1) attrition of prosody, and little research has investigated the effect of individual differences in late bilinguals in terms of L2 acquisition and L1 attrition of prosody. In the current study, I therefore investigate L2 prosodic acquisition and L1 prosodic attrition in late E-A and A-E bilinguals in the production of their *wh*-words. I also explore whether late L2 learners' language use and phonetic aptitude affect how target-like learners are in the prosody of both English and Arabic *wh*-words. Fifty-nine participants participated in the study: 14 E-A bilinguals, 15 A-E bilinguals, 15 monolingual English speakers and 15 monolingual Arabic speakers. Both bilingual groups had been living in the L2 environment for about 20 years. All participants read brief dialogues of *wh*-question/answer pairs, and the bilinguals read them in both Arabic and English. The bilinguals also completed a language-background questionnaire, a proficiency test and a phonetic aptitude (talent) test. The pitch contour of the *wh*-question words was analysed using growth-curve analyses to capture differences in slope steepness and amount of curvature. The results from the monolingual groups revealed a steep rise to a high target on Arabic *wh*-words, but no such high target in English *wh*-words. The A-E and E-A bilinguals approximated the prosodic patterns of the L2. In addition, the results revealed an asymmetrical pattern of L1 attrition, showing attrition among the E-A bilinguals, but not among the A-E bilinguals. Language use, but not (or at least to a much lesser extent) phonetic aptitude, modulates how closely bilingual participants approximate native patterns.

Keywords

Bilinguals, second-language acquisition (L2), first-language (L1) attrition, *wh*-words, language use, phonetic aptitude (talent), pitch contour

2.1 Introduction

Late bilinguals who speak a second language (L2) frequently encounter challenges in perceiving or producing segments that either differ from or do not exist in their native language (L1). Previous studies have reported on these challenges in detail, and various theoretical models have been developed to account for inaccurate production and perception of the phonetic segments in the L2 (Flege, 1986, 1991, 1995; Best et al.,1988; Kuhl, 1991; Kuhl, et al., 1992; Best, 1994, 1995; Iverson & Kuhl, 1995, 1996; Major, 2001). In addition, L2 learners are known to make prosodic and intonational errors (Willems, 1982; Flege, 1992, 1995; Flege et al., 1995; Munro & Derwing, 1995; Magen, 1998). However, advanced knowledge of an L2 may also influence a learner's L1. In particular, late L2 learners sometimes differ in the production of their L1 compared to other native speakers of their L1 (Mennen, 2004); this phenomenon is known as L1 attrition (Freed, 1982; Schmid, 2004). De Leeuw (2008, p.10) defined L1 attrition as the "non-pathological, non age-related structural loss of a first language within a consecutive bilingual", i.e. a bilingual who has learned the L2 after the age of three (Baker, 2011). This definition excludes language loss due to injury or illness, but it includes loss through maturation that is not attributable to the aging process (Goral, 2004). Thus, when individuals do not actively use their L1 over a long period, or if they use their L2 more than their L1, their L1 might differ from that of monolingual native speakers due to L2 influence on the L1. When L1 attrition influences the domains of phonetics or phonology at the segmental or suprasegmental level, speakers would be perceived as having a foreign accent in their L1 (de Leeuw, 2010; Hopp & Schmid, 2013; Bergmann et al., 2016).

Numerous factors play a role in achieving or maintaining native-like speech. These include the daily linguistic environment of a bilingual and the quantity and quality of L1 and L2 exposure (de Leeuw, 2010; Mayr et al., 2012). Furthermore, phonetic aptitude (i.e. talent) plays a role in both L1 attrition and L2 acquisition (Jilka et al., 2008; Hopp & Schmid, 2013). The present study focuses on the pronunciation, particularly the intonation pattern of *wh*-words in English and Arabic. I investigate the intonation of *wh*-words by monolingual native speakers of both languages and the relationship between L1 attrition and L2 acquisition in highly fluent, late consecutive A-E and E-A bilinguals. Furthermore, I explore how the amount of L2 language use and phonetic aptitude influence both L1 attrition and L2 acquisition of intonation (see chapter 1, section 1.5, for why these particular factors were selected).

2.2 Literature Review

2.2.1 L2 Acquisition of Prosody

Over the few past decades, most L2 acquisition research has focused on segmental information, such as VOT or vowel formants (Weinreich 1954; Flege, 1994, 1995; Chang, 2012, de Leeuw et al., 2013), and fewer studies have investigated the suprasegmental aspects of L2 acquisition (Willems, 1982; McGory, 1997; Ueyama, 1997; Mennen, 1999; Guion et al., 2000; Jilka, 2000; de Leeuw et al., 2013). Consequently, knowledge about aspects of L2 prosodic acquisition in late consecutive L2 learners is limited.

Similar to studies on the L2 acquisition of segments (cf. Eckman, 1981; Flege, 1986, 1991, 1995; Best, 1994, 1995; Major, 1992, 2001), several studies have reported the L1 transfer of prosodic properties into the L2, even in highly advanced learners (Willems, 1982; Elst & de Bot, 1987; Ueyama & Jun, 1996; McGory, 1997; Mennen, 1999; Ramirez & Romero, 2005; Horgues, 2010). For example, Mennen (1999) investigated peak alignment in pre-nuclear pitch accents in highly proficient native Dutch L2 learners of Modern Greek, and she found that the peak is aligned earlier in Dutch than in Greek. The results concerning highly proficient learners showed that they aligned the peak in their L2 Greek as early as they would in their L1 Dutch, thereby suggesting L1 transfer into their L2 Greek.

Ramirez and Romero (2005) studied the intonational patterns of tag questions such as *It's cold today, isn't it?* in native English speakers and native Spanish L2 learners of English. They found that the non-native English speakers used a rising pitch contour to confirm the information in the tag questions, while the native English speakers used a falling pitch in this context. Ramirez and Romero (2005) suggested that L2 speakers' rise in pitch in tag questions is due to L1 transfer because Spanish tag questions are characterised by a rising pitch. Similarly, Willems (1982) addressed the impact of the L1 on the L2 by examining the intonation patterns of native Dutch speakers learning English in terms of both their production and perception of English. He found that the participants' productions in English (L2) were different from monolinguals (native English), which he suggested was a result of the influence of their native language (Dutch). Deviations from the monolingual English norm included a narrower pitch range than is typical for English, rises when falls were expected and vice versa, unstressed syllables with uncharacteristically high pitch and final rises that are too low in pitch.

By contrast, Ueyama and Jun (1996) found that some intonational errors cannot be traced back to speakers' particular L1, but may be common in L2 learners of English across different L1 backgrounds. They analysed focus realisation in interrogative and declarative sentences produced by Seoul Korean and Tokyo Japanese speakers who were L2 learners of English and at different levels of proficiency. While the prosodic systems of the three languages differ, focus is realised quite similarly. For example, both English and Korean are characterised by having no pitch accents after the focused element and preceding the following boundary tone. Japanese also often shows this kind of dephrasing after the focused element, but sometimes realises an H tone on the word following the focused element. Despite these similarities across the three languages in terms of dephrasing, Ueyama and Jun (1996) found an inverse relationship between learners' proficiency and the number of pitch accents after the focused element. In other words, more advanced learners showed more dephrasing and thus more native-like productions than did less advanced learners, regardless of L1 background and despite dephrasing being common in both learner groups' L1s.

To summarise, the selection of studies presented in this section provide evidence that transfer or interference from the L1 plays a role in L2 speakers' productions of intonation, but that not all deviations from the native-speaker norms can be explained in terms of L1 transfer or interference.

2.2.2 L1 Attrition of Prosody

Much research on L2 acquisition is concerned with how the L1 inhibits, interferes with or transfers to L2 acquisition (Gass, 1983). In addition, much of the empirical work focuses on understanding how people can be efficiently and effectively supported in acquiring an L2. However, research has shown that an L2 can also inhibit, interfere with or transfer to the L1, thus, giving evidence for bidirectional influence. Bilinguals whose dominant language is the L2 are more likely to show L1 attrition than those whose dominant language is the L1 (Schmid & Köpke, 2007). Some studies have focused on how the L2 influences L1 segments (Flege & Hillenbrand, 1984; Flege, 1987a, 1987b; Major, 1992; Chang, 2012; Mayr et al., 2012; de Leeuw et al., 2013; Bergmann et al., 2016), while others have focused on how the L2 influences L1 prosody (Mennen, 2004; de Leeuw, 2009; de Leeuw et al., 2012).

Mennen (2004) investigated the influence of the intonational acquisition of L2 on learners' L1. She studied the temporal alignment of pitch peaks in fluent Greek-Dutch and Dutch-Greek bilinguals and monolingual controls. The phonological structure of Dutch and Greek is the same in non-final and pre-nuclear rises; however, the two languages differ in the timing of these rises with an earlier peak in Dutch. Furthermore, the length of the vowel in the accented syllable affects the timing of the peak in Dutch with earlier peaks for longer vowels and later peaks for shorter vowels, but not in Greek. Mennen (2004) found that the native Dutch speakers showed bi-directional interference when producing the pre-nuclear rising accent. This interference was particularly prominent in four of the five speakers involved, whose peak alignment differed from that of the native control group of both languages. However, peak alignment similar to native speakers was produced by the fifth speaker in both languages.

Similarly, de Leeuw (2009) assessed L1 attrition with regard to the attainment of pre-nuclear tonal alignment in German and English. Tonal alignment is similar in German and English in that the peak and the fall-rise in nuclear position occur late in both languages (Ladd, 1996). However, tonal alignment in pre-nuclear syllables differs across the two languages with an earlier rise in English compared to German and an earlier rise in Northern German compared to Southern German. de Leeuw (2009) examined monolingual English speakers, monolingual German speakers and native German L2 English learners living in Anglophone Canada. The timing of the beginning of the rise in pre-nuclear syllables occurred earlier in the bilingual Germans than it did in the monolingual German controls, suggesting L1 attrition due to the influence of English.

While there is increasing evidence that a speaker's L2 can influence his or her L1, researchers have proposed several different mechanisms for how the L2 influences the L1. De Leeuw et al. (2012) suggested that L1 attrition may occur because speakers already have gaps in their native language and that these gaps are filled with components of the L2. Furthermore, de Leeuw et al. (2012) also suggested that L1 attrition might be due to the dynamic growth of the L1, a development that extends further than the subsystems of the L1 and L2. Alternatively, Schmid (2002) proposed that attrition occurs because the L2 elements interfere with the L1.

Hopp and Schmid (2014) investigated L1 attrition and L2 acquisition through foreign accent ratings. Their approach differs from the previously mentioned studies in that it does not focus on a particular aspect of prosodic structure, but rather represents a more global measure of phonetic and phonological attrition. In particular, Hopp and Schmid (2013) compared L1 German attriters (selected from de Leeuw et al., 2010) and late L2 learners of German with Dutch or English as their L1 (selected from Hopp, 2007) with German monolinguals. Native German speakers judged the extent to which the speakers had a foreign accent in their German. As expected, the late L2 learners of German had a stronger foreign accent at the group level than both groups of L1 German speakers, which was unrelated to the learners' length of residence (LOR). Moreover, Hopp and Schmid (2010) found a sizeable overlap between the L1 attriters and the L2 learners: about 80% of learners scored within the L1 attriter range. Hopp and Schmid (2013, p.388) concluded "that acquiring a language from birth is not sufficient for ensuring nativelikeness in bilingual speech production".

In summary, several studies suggest a bi-directional influence of L1 and L2 prosody that is measurable at a fine level of detail where languages differ in their prosody and that manifests in more global measures, such as bilinguals' perceived foreign accent.

2.2.3 Factors that Influence L1 Attrition and L2 Acquisition

Individuals differ from each other in their L2 acquisition ability and the extent of their L1 attrition. Köpke and Schmid (2004, p.15) listed "age at onset of L2 acquisition, age at onset of L1 attrition, time since onset of attrition, level of education, attitudes, frequency, amount and settings of use of the attriting language" as factors that may affect L2 acquisition ability and the degree of L1 attrition. Many previous studies on L2 acquisition or L1 attrition have focused on factors that may affect performance in the phonetic domain. These include age of arrival (AOA) or LOR (e.g. Flege, 1987a, 1987b; Flege et al., 1995; Bongaerts et al., 1995; 1999; Mennen, 1999; Abu-Rabia & Kehat, 2004; Rasier & Hiligsman, 2007; see Piske et al., 2001, for an overview), age of learning (e.g. Johnson & Newport, 1989; Hopp & Schmid, 2013) and L1/L2 frequency of use (e.g. Piske et al., 2002; de Leeuw, Schmid &

Mennen, 2010). Of relevance for the current study are the amount of L1/L2 language use and phonetic aptitude (i.e. talent).

The Role of Language Use in L2 Acquisition and L1 Attrition

Previous studies have examined the influence of linguistic experience on L2 acquisition, which is often operationalised as the LOR in an L2 country or the amount of L1 and L2 use (Trofimovich & Baker, 2006). The impact of language use and LOR on L2 acquisition and L1 attrition has been examined in relation to learners' knowledge of segmentals and suprasegmentals (Purcell & Suter, 1980; Leather, 1987; Shen, 1990; Thompson, 1991; Flege & Fletcher, 1992; Moyer, 1999; Guion et al., 2000; de Leeuw, 2009; Mayr et al., 2012). However, few studies have looked specifically at everyday L1 and L2 language use or the role of the quantity and quality of language use when maintaining native-like L1 speech patterns after moving to an L2 environment (de Leeuw et al., 2009; Mayr et al., 2012).

Research has shown that the AOA in migrant populations influences the amount of L1 and L2 language use. Early L2 learners (with ages of arrival younger than 10) seem to use the L2 as the primary means of communication more often than they use the L1. However, late L2 learners use their L1 more frequently when talking to members of their home or native community and in the target-language environment (cf. Piske et al., 2001; Jia, Aaronson & Wu, 2002). AOA seems to not only influence the amount of L1 and L2 language use, but also affect L2 acquisition and L1 attrition, with more native-like L2 productions and more L1 attrition in earlier arrivals compared to later arrivals (Flege et al., 1995; Flege et al., 1999; Lee, 2011).

However, most other factors related to the amount of L2/L1 language use seem to have a small impact on L2 acquisition. This finding may be because the amount of language use in an average day or week can be more difficult to measure objectively than the AOA. Subjective measures (such as learners' statements and reports) are often used to determine how often bilinguals use their L1 and L2 on a daily basis. In a meta-analysis of data from 240 L1 Korean and 240 L1 Italian L2 learners of English across several tasks and studies, Flege (2009) reported that the quantity of L1/L2 use accounted for less than 10% of the difference in foreign accent rating data. This low influence may be caused by inaccuracies in participants' reporting of L1 and L2 language use. Alternatively, it may not be the quantity but the quality of L1 and L2 use that has a greater effect on L2 acquisition and L1 attrition. For example, neurobiological research on language learning and stimulus appraisal has implied that emotional involvement might be necessary for effective L2 learning (Schumann, 1998). This idea is further supported by Hopp and Schmid's (2010) study, which found that the more often L2 learners of German could engage in conversations with native German speakers in German, the better was their L2 acquisition. Similarly, Suter's (1976) and Purcell and Suter's (1980) early research into the influence of language use on the degree of a foreign accent in L2 learners supports this idea. In their studies' questionnaires, the learners of English estimated how much they used English at work, home and in school with native speakers of English and how much time they spent with English native speakers. Purcell and Suter (1980) identified cohabitation with English native speakers (i.e. more L2 use in the home) as being the third most important variable affecting the degree of a foreign accent in L2 learners. By contrast, owing to its correlation with L1 background and LOR, the amount of L2 use at work or school was not a significant factor affecting the degree of an L2 foreign accent.

While the previous studies have revealed a small influence of language use on L2 acquisition, other studies have revealed no obvious impact from L2 use or input. For example, Flege and Fletcher (1992) found no significant correlation between the percentage of everyday L2 English use and the degree of an L2 foreign accent for their native Spanish-speaking participants. Thompson (1991) asked L1 Russian immigrants to estimate how often they used English at work, at home and with friends. She found a simple and significant correlation between the amount of English use and the degree of an L2 foreign accent. However, an analysis of multiple regression did not confirm the amount of English use as being a significant predictor (since it was confounded with age of learning). She also hypothesised that the L1 has a possible effect on L2 pronunciation since some immigrants with high L1 proficiency had an L2 foreign accent. Based on her hypothesis, she argued that "a difference must be noted between subjects who have maintained their mother tongue and those who have lost it when it comes to estimating accent retention in the second language" (Thompson, 1991, p.200).

Previous studies have also examined the impact of language use on L1 attrition (Schmid & Köpke, 2007; de Leeuw et al., 2009). L1 contact seems to play an important role in maintaining an L1 accent. de Leeuw et al. (2010) sought to determine whether native German speakers residing in the Netherlands or (Anglophone) Canada are perceived as having a foreign accent in their L1 German. Nineteen German monolingual listeners assessed the foreign accents of five German monolinguals and 34 L1 German speakers residing in Canada and 23 L1 German speakers living in the Netherlands, who had moved to the host country at an average age of 27 and who had been living there for an average of 37 years. The results showed that nine of the immigrants to Canada and five of the immigrants to the Netherlands were perceived as being non-native German speakers. Importantly, de Leeuw et al. (2010) found that the quality and quantity of language contact with German native speakers had a greater influence on predicting a global foreign accent in native speech than the LOR or AOA. In particular, they found that the amount of L1 use in communicative settings where code-switching was unlikely to occur predicted foreign accent ratings of the L1 speech, but the amount of L1 use in communicative settings where code-switching was likely to occur did not. Thus, speaking the L1 in settings where code-switching is unlikely to occur seems to be conducive to maintaining a high level of native language pronunciation within consecutive bilinguals in a migrant context.

Overall, it appears that certain kinds of language use, such as that used in the home or in situations where code-switching is unlikely, have an influence on L1 attrition. Notably, as most of these studies concern a global foreign accent, it is unclear which aspects of pronunciation may be affected by language use. To contribute to a more detailed picture of the role of language use in L2 acquisition and L1 attrition in the phonetic domain, the current study focuses on one particular aspect of question intonation. Thus, the objective of the present study is to offer a narrower, but more in-depth investigation of the effect of language use on L2 prosody learning and L1 attrition.

Role of Aptitude in L2 Acquisition and L1 Attrition

Carroll (1993) defined foreign language aptitude (or talent) as a cognitive ability that helps a person master an L2/foreign language. In contrast, phonetic talent refers to a person's ability to master the pronunciation of the foreign/second language (cf. Jilka et al., 2011). Within a normal population, the degree of language aptitude has been shown to vary substantially, unrelated to other individual factors (Novoa et al., 1988; Schneiderman & Desmarais, 1988; Ross et al., 2002; Skehan, 2002; Dörnyei & Skehan, 2003). According to

Bylund et al. (2009, p.447), one of the most important elements of language aptitude is "phonetic/phonemic coding ability, that is, the capacity to identify speech sounds and to make sound-symbol associations" (for more information, see Carroll & Sapon, 1959; Carroll, 1981). Jilka et al. (2008, p.171) defined phonetic talent as "an innate, somewhat mysterious ability that a person either has or does not have".

There is general agreement that there is a practical difference between proficiency (the observed, perceived ability of someone to perform a particular task) and talent, despite the two often being confused. One might be extremely proficient to perform a particular task, having practiced for many years, without exceptional aptitude. Therefore, talent is considered an inherent characteristic, which cannot be learned.

Some have sought to examine the brain to see if there are any physiological indicators of talent for any given skill. Diamond et al. (1985), for example, attempted to look for such clues in the brain of Albert Einstein. Studies of this kind, however, are not widespread. Others, such as Geschwind and Galaburda (1985), have claimed to find a link between the increased growth of certain areas of the brain and a 'pathological language talent'. Important neuropsychological models, such as those of Gardner (1983), which have sought to 'locate' talent in the brain, have found it 'next to' other special abilities, such as musical, spatial and logical talent.

In most instances, when language 'talent' is measured, it is done so by testing general ability. However, there have been some tests devised to measure the exact nature of 'pathological', or exceptional ability to learn languages. One example of such a test is Novoa et al.'s (1998), which compared ability to "IQ, vocabulary skill, general language aptitude, verbal fluency, verbal memory, apprehension of abstract patterns, understanding (digit symbol test), and learning of code systems" (Jilka et al., 2008, p. 244)

It is assumed that language 'talent' is made up of many different linguistic skills. Schneiderman and Desmarais (1988) consider talent for accent to be discreet from talent for grammar, for example. This is illustrated quite well by the "Joseph Conrad Phenomenon" (Bongaerts et al., 1995; Bongaerts, 1999; Guiora, 1990), which tells of the Polish writer's exceptional English grammar and vocabulary abilities, whilst his accent was still very strongly Polish in pronunciation.

Phonetic aptitude has received less attention in the L2 acquisition and L1 attrition literature than some of the other variables such as LOR and age of acquisition. The few previous studies on aptitude have focused on L2/foreign language learning (Tahat et al., 1981; Flege et al., 1995; Thompson, 1995), with the exception of a few studies by Skehan (1989) and Skehan and Ducroquet (1988) that examined the role of aptitude on L1 development in a monolingual L1 setting. To the best of my knowledge, there is no study on the role of aptitude on L1 attrition that uses specific aptitude tests to measure the participants' aptitude.

Several studies have used self-ratings or indirect measures of phonetic aptitude to investigate how talent influences L2 acquisition, especially L2 pronunciation. For example, Hopp and Schmid (2013) used a proficiency test as an indirect measure of aptitude and suggested that aptitude is linked to the presence of a foreign accent for the bilinguals in their study. Neufeld (1979) suggested that a relationship may exist between a foreign accent and musical ability, such as the ability to distinguish musical pitches. However, while Suter (1976) did not confirm such a relationship, Flege et al.'s (1999) results suggest that musical ability may be involved in bilinguals' degree of foreign accent. Participants in Suter's (1976) study mimicked sections of speech containing unfamiliar sounds, and no significant effect of aptitude (musical ability) was found on the degree of the L2 foreign accent. Flege et al. (1999) found that a measure combining self-reported mimicry and musical ability with the ability to remember how L2 words are pronounced accounted for a small percentage of the variance in L2 foreign accent ratings. However, these studies all used self-ratings rather than phonetic aptitude tests. Other studies have used mimicry ability as a measure of phonetic talent. For example, participants in Flege et al. (1995) repeated tape-recorded sentences, and the results showed that mimicry ability played a significant role in the degree of the L2 foreign accent (cf. also Li, 2016).

Several authors have advocated the use of aptitude tests that measure language aptitude directly in place of self-ratings or indirect measures of aptitude (cf. Jilka et al., 2010; Hinton, 2012). The Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959) is made up of five sections, each section is designed to gauge a certain skill or ability pertaining to foreign language learning. The first section is called Number Learning. This section sees participants learn a series of numbers through aural input, and then differentiate between a series of different combinations of those numbers. Secondly, in the Phonetic Script section, examinees are asked to learn the relationship between a set of speech sounds and phonetic symbols. In the third section, called Spelling Clues, participants read words, which have been spelt phonetically (as they are pronounced), rather than how they normally would be spelt. Then, they select a synonym of the "disguised" word from a list. In section four, Words in Sentences, their awareness of grammatical structure is tested. The participants are presented with a sentence and a key word is highlighted in the sentence, which performs a specific role. Then, they are shown a second sentence (or sentences) and asked to highlight a word, which performs the same role as the key word in the first sentence. Lastly, the Paired Associates section sees participants learn a series of foreign language words, then memorise their definitions in English.

Jilka et al. (2008) conducted a large-scale project to examine pronunciation talent of L2 speakers. They used three tasks of the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959) test to measure language aptitude directly: The 'Spelling Cues' test, which explores participants' ability to associate symbols with sounds, the 'Words in Sentences' test, which measures participants' grammatical awareness and the 'Paired Associates' test, which measures participants' ability of rote memorisation (Jilka et al., 2008). They tested 105 German native speakers with L2 English and 15 native English speakers in English and found that phonetic talent correlates with how well the speakers performed the spoken tasks (Jilka et al., 2008).

Finally, Mennen's (2004) study sheds some light on how phonetic talent may be related to native-likeness in both the L1 and the L2. Her work highlighted the fact that some participants can be seen as 'outliers' from the main trend because they can maintain native-like abilities in their L1 and their L2, which raises the question: What makes these individuals particularly good at learning the L2 while maintaining nativeness in their L1? One possibility is that participants who do well in the L2, and who also show no effects on the L1, may be phonetically talented, which could explain their success. The present study will use specific aptitude tests to examine the effect of phonetic aptitude on the prosody of *wh*-words in the L1 and the L2.

2.2.4 Prosody of *Wh*-questions in Modern Standard Arabic (MSA) and English

MSA is a standardised form of Arabic that has been adopted as a *lingua franca* and can be understood in spoken and written form by speakers of Arabic, regardless of their local dialect. MSA is used in official communications and in the media, such as newspapers and television; it is also taught in schools and universities. Since MSA is codified in grammar books and dictionaries, it varies little across the Arabic-speaking world (Alosh, 2005). However, greater variety exists in regional forms of Arabic and in socially distinct forms, such as formal and colloquial registers. Like Arabic, English is a pluricentric language with spoken and written forms that vary across different regional areas and social groups, and that has several well-known forms, such as British English and American English (Clyne, 2012).

While researchers have widely discussed the syntax of *wh*-questions in Arabic, English and other languages (Chomsky, 1977, 1995, Aoun & Benmamoun, 1998; Khomitsevich, 2008), the prosody of the *wh*-word in *wh*-questions remains largely unexplored, especially in Arabic. To date, few studies have been conducted on the prosody of *wh*-words in *wh*-questions in English, Arabic and other languages. Most studies that have been conducted on the prosody of *wh*-questions have focused on utterance-final intonation (Pierrehumbert & Hirshberg, 1990; Celce-Murcia et al., 1996; Syrdal & Jilka, 2003; Alotaibi 2013; Hellmuth & Almbark, 2015). To the best of my knowledge, the present study is the first large-scale investigation focusing on the prosody of *wh*-words in MSA and English.

In MSA, a wh-question is formed with an interrogative marker (IM). For the purposes of the current study, I refer to IMs in MSA as wh-words because they are analogous to English *wh*-words. The MSA *wh*-words used in this study are **من** (/man/, Engl.: *who*), **وماذا** (/mɑːðɑː/, Engl.: *what*), أين (/ajna/, Engl.: *where*) and متى (/matɑː/, Engl.: *when*). In MSA, the *wh*-word in *wh*-questions occurs at the beginning of the *wh*-question, is typically stressed and is typically characterised by a high rise in pitch (Defense Language Institute, 1974). Wh-questions in MSA occur with the typical Verb–Subject–Object (VSO) word order, and not with the more marked Subject–Verb–Object (SVO) word order (cf. Fakih, 2015). Yes–no questions in MSA do not contain an IM and differ from declaratives in their intonation (Defense Language Institute, 1974). Examples (1) and (2) show an example of a *wh*-question with a stylised intonation contour in English and Arabic, respectively.

(1) And when are you going?

(2) Wa mata sawfa naðhab?

And when will$_{AUX}$ go$_{SG/FUT}$

"And when are you going?"

El Zarka's (1997) corpus study of television interviews looked at *wh*-questions in the MSA of Egyptian Arabic speakers and reported that *wh*-questions in MSA are characterised by an early peak in the question and a lower final pitch range at the end of the *wh*-questions. She also reported that the *wh*-word at the start of the *wh*-question has a higher pitch level when it is the initial word in an MSA sentence and that it is always accented (El Zarka, 1997). This confirms the previous observations of a high rise in pitch on the *wh*-word. In particular, El Zarka (1997) suggested that the pitch accent on the *wh*-word in Arabic is a ↑H* (L), where the ↑ in indicates a raised fundamental frequency (F0), which means is that there is a high peak on the wh-word with a following fall.

Only a few studies have investigated the pitch accents on wh-words in English (Steedman, 2000; Hedberg & Sosa, 2002; Bartels, 2014). Bartels' (2014) theoretical work assumes that the *wh*-word in English typically receives an H* accent or no pitch accent at all. Lambrecht and Michaelis (1998, p.515) proposed that the *wh*-element is not likely to be accented because "there is no commitment on the part of the speaker to the effect that she knows the identity of the referent of the expression". In particular, the *wh*-word is not accented for the same reason that indefinite expressions, such as *something* or *someone*, are not accented in declarative sentences: "There simply is no referent that the addressee is expected to identify at the time of utterance" (Michaelis, 1998, p.515).

Although some previous studies concentrated on the nuclear contour of *wh*-questions in English, but they also provided brief information about the prosody of the *wh*-word (Steedman, 2000; Hedberg & Sosa; 2002). Steedman (2000, p.1) suggested that the *wh*-word is "the 'theme' of the wh-question because it evokes but does not select from an alternative set". Therefore, he proposed that the English *wh*-word would be marked with an L*+H or L+H* pitch accent (Steedman, 2000). Hedberg and Sosa (2002) studied a corpus of 73 negative and positive yes-no questions and *wh*-questions to examine both the *wh*-word and question-final intonation of English *wh*-questions. They found an L+H* pitch accent in most of the *wh*-words in *wh*-questions and the initial auxiliary in yes–no questions (Hedberg &

Sosa, 2002). In particular, they determined that 60% of *wh*-words were marked with an L+H* and 26% were marked with an H* pitch accent (Hedberg & Sosa, 2002).

To summarise, the few studies that have examined the prosody of *wh*-words in *wh*-questions in MSA and English suggest a steep rise in pitch on the *wh*-word in MSA. However, the findings for English are inconsistent.

2.3 The Present Study

The main aim of this study is to examine whether highly-fluent late consecutive A-E bilinguals who moved to the United Kingdom (UK) in adulthood and highly-fluent late consecutive E-A bilinguals who moved to the Kingdom of Saudi Arabia (KSA) or Yemen in adulthood show native-like prosody in their L2 and/or show L1 attrition of prosody. Arabic and English were chosen because they are typologically unrelated. I thus expand on previous work on L1 attrition, which has mostly focused on Germanic languages, and I compare two typologically different languages. In the current study, the intonation of the *wh*-word in *wh*-questions is the phonetic variable selected because there are indications in the literature that the intonation of the *wh*-word in *wh*-questions may differ in English and Arabic. Additionally, few studies on the intonation of *wh*-questions have focused on the *wh*-word. Since quite little is known about the prosody of the *wh*-word in Arabic and English *wh*-questions and since the dataset for the current study is large, I rely on the acoustic signal for my analyses rather than manual intonational annotation, such as ToBI annotations (Silverman et al., 1992). Many previous studies doing so have looked at pitch peak alignment. However, there is no indication in the previous literature that Arabic and English *wh*-words may differ in their alignment of pitch peaks. Thus, I may find a more pronounced pitch excursion in Arabic compared to English since a steep rise in pitch has been reported for Arabic *wh*-words, but not consistently for English. I therefore need to conduct an analysis that captures peak height as well as the presence and absence of a peak. In the current chapter, I use growth curve analysis (GCA; Mirman, 2014; Winter & Wieling, 2016) to capture differences in the curvature of the pitch contour across the *wh*-word. This analysis requires manual annotation of the beginning and end of the wh-word only and can be conducted even on unaccented wh-words.

The second aim is to determine whether language use has an effect on L1 attrition and L2 acquisition of prosody. Language use was selected as a factor of investigation because it seems to be especially relevant for L1 attrition because many researchers assume that prolonged use and exposure to the L2 is needed for speakers' L1 pronunciation to attrite. All of the current study participants are late L2 learners who have lived in the L2 environment for a substantial amount of time and thus have extensive exposure to the L2. I therefore measure language use by means of a detailed questionnaire that gauged how often participants use each language and with whom.

The final aim of this study is to ascertain whether aptitude plays a role in L2 language acquisition and L1 attrition of prosody. To do so, I apply some phonetic aptitude tests developed by Jilka (2007, 2009) and Jilka et al. (2008, 2011) to measure aptitude directly. Phonetic aptitude was selected as the second factor of investigation because its relationship to L2 acquisition and L1 attrition is unclear. In particular, it may be that participants with high phonetic aptitude show more native-like pronunciation in their L2, but that this happens to the detriment of pronunciation in the L1. Alternatively, it may well be that learners with a specific aptitude for the phonetic aspects of language are better at inhibiting L1 attrition, while at the same time outperforming those with less phonetic aptitude in the L2. That is, those learners with 'a good ear' may be better at maintaining a native accent in their L1 while also achieving more native-like L2 productions (cf. Mennen, 2004).

2.4 Hypotheses

This study explores L1 prosodic attrition and L2 prosodic acquisition in the production of *wh*-words in late E-A and A-E bilinguals, i.e. bilinguals who speak two typologically different languages. This study also explores whether late L2 learners' phonetic aptitude and language use influence how target-like they are in the prosody of both Arabic and English *wh*-words. The current study explores the following hypotheses:

1. *Arabic and English patterns:* Based on the previous literature, I would tentatively expect a difference in the pronunciation of *wh*-words in Arabic and English, with a steeper rise in pitch in the *wh*-word in Arabic than in English.
2. *L2 acquisition:* Based on LILt, I would expect that the highly proficient E-A and A-E bilinguals will approximate the L1 pattern in their L2.

3. *L1 attrition:* Based on LILt and since other studies have found L1 attrition in people who have been in the host country for a significant amount of time, I expect that – as with other aspects of one's L1 – prosody will show attrition. Thus, I expect A-E bilinguals to diverge from the native norm in the direction of their L2 English (i.e. a shallower rise in pitch on the *wh*-word than English monolingual controls) and the E-A bilinguals to diverge from the native norm in the direction of their L2 Arabic (i.e. a steeper rise in pitch on the *wh*-word than Arabic monolingual controls).
4. *Language use:* Based on the existing literature, I expect a small influence of language use on L1 attrition and L2 acquisition, such that bilinguals with more L2 use will show a more native-like pattern in their L2 compared to those with less L2 use.
5. *Phonetic aptitude:* While there is no consensus in the literature regarding the role of phonetic aptitude in L2 acquisition and L1 attrition, I tentatively assume that participants with high phonetic aptitude will show more native-like production of *wh*-words in both their L1 and L2. This hypothesis is based on the observation that participants in Mennen's (2004) study showed native-like productions in either both their languages or in neither.

2.5 Methods

2.5.1 Participants

The study cohort consisted of 59 participants divided into four groups. Fifteen of the 59 participants were monolingual native Arabic speakers from different regions in Saudi Arabia: Makkah, Jeddah, Riyadh and Abha (four males, 11 females; mean age = 36.06; SD = 6.91), and 15 of the participants were monolingual native English speakers from different regions of the UK: Sheffield, Chester and London (two males, 13 females; mean age = 40.93; SD = 9.65). These 30 participants, which comprised the Arabic and English control groups (Monolingual Arabic Group and Monolingual English Group, respectively), were selected through the snowball sampling strategy, with personal contacts in the UK and Saudi Arabia as initial participants, and where each participant then provides the name of another potential participant. This means that some recruitment bias may arise in this sample due to the connectedness of their selection.

The two monolingual groups were matched with the bilingual groups in terms of regional dialect as closely as possible. The A-E bilingual group included 15 participants (one male, 14 females; mean age = 39.4; SD = 3.75) who were native MSA speakers from Saudi (Jeddah, Makkah or Riyadh) or from Yamani (from eastern Yamani), but had lived most of their lives in Makkah and Jeddah. Both regional dialects (Saudi and Yamani Arabic) are close to each other and to MSA. The A-E bilinguals were also highly proficient in their L2 English. To be included in the study, A-E bilingual participants had to have fully acquired their L1 in their home country, moved to the UK after the onset of puberty (mean AOA = 19; SD = 2.8) and had to have lived in the UK (Sheffield, Chester and London) or 15 years or longer at the time of the study (mean LOR = 20 years; SD = 2.6). Some of the A-E bilinguals had also lived in the United States (US) prior to relocating to the UK, but none had moved there before the onset of puberty. As with the monolingual Arabic and monolingual English groups, the participants in A-E bilingual group were selected using snowball sampling. Two additional A-E bilinguals were excluded from the study due to low results in the proficiency test.

The E-A bilingual group included 14 native English speakers with a high proficiency in MSA (four males, 10 females; mean age = 34.53; SD = 3.6). Again, to be included in the study, the E-A bilinguals had to have fully acquired their L1 in their home country, in this case in Sheffield or London, moved to KSA (Makkah and Jeddah) or Yemen after the onset of puberty (mean AOA = 18; SD = 1.60) and had to have lived in KSA or Yemen for 15 years or longer at the time of the study (mean LOR = 19; SD = 3). Some of the E-A bilinguals who had been learning their L2 for several years in Yemen had continued their learning in the UK at a regular Arabic school or mosque for two to three years and then returned to KSA or Yemen. Several of the participants were in the UK at the time of the study for a holiday and the rest were in KSA at the time of the study. The participants were recruited by advertising through the Arabic Language Institute in Umm Al-Qura University (ALI, UQU; see Appendix I for approval consent from ALI, UQU) and by applying the snowball sampling strategy. Sixteen additional E-A bilinguals were excluded from the study; two of the participants had knowledge of Urdu (which was an exclusion criterion because Urdu was part of the aptitude test I used in this study) and the others were not highly proficient in their L2 according to a proficiency test.

2.5.2 Materials and Procedure

Production of *wh*-Questions

In this study, a reading task was used to elicit the pitch accent patterns of Arabic and English *wh*-questions. Participants were asked to read a list of 12 *wh*-questions and eight yes–no questions (as filler sentences) six times in both languages, so that each participant produced 72 *wh*-questions in each language. The *wh*-words in Arabic were either monosyllabic (/man/=من , Engl.: *who*) or bisyllabic (/mɑːðɑː/ =وماذا, Engl.: *what*; /ajna/= أين , Engl.: *where*; /matɑː/= متى Engl.: *when*) while the English *wh*-words were all monosyllabic (*who*, *what*, *where*, *when*). Since Arabic has only one monosyllabic *wh-word*, and the other *wh-words* in Arabic are bysilabic, I have included both monosyllabic and bisyllabic Arabic *wh-words*. In contrast, English only has monosyllabic *wh-words*. All *wh*-questions were phrased such that the *wh*-word was the second word in the sentence. This word order was used so that the pitch curves were not influenced by any possible initial boundary tones (Arvaniti et al., 1998). Yes–no filler questions were included to prevent participants from using list-intonation. The stimuli sentences were first created in English and then translated into Arabic. A native speaker of each language checked the sentences to make sure that they sounded natural and contained no mistakes. Participants were told to read the questions as naturally as possible, as if they were seeking information in a neutral context. To further encourage readings that were as natural as possible and to elicit pitch accent patterns that resembled a natural dialogic situation (Jilka et al., 2008), the sentences were integrated into mini-dialogues, as shown in examples (3) and (4) for English and examples (5) and (6) for Arabic (with target *wh*-words in boldface; for more examples see Appendix C):

(3) *From* ***where*** *is your friend Linda leaving?*
From London.

(4) *And* ***what*** *did you do in London?*
We spent three wonderful weeks there.

(5) wa **mɑːðɑː** faʕalt fiː landan
and what do$_{SG/PST}$ in London
And what did you do in London?
qadˤinɑː θalɑːθ ɑːsabiʕ ʒamiːlah
spend$_{PL/PST}$ three week$_{PL}$ wonderful
We spent three wonderful weeks there.

(6) wa **matɑ:** saufa nahbIṭˤ?
and when will$_{AUX}$ go$_{SG/FUT}$
And when are we landing?
lɑ: ɑ:ʕlam
NEG know$_{SG}$
I don't know.

Background Questionnaire

The purpose of the questionnaire was to obtain sociolinguistic background information for the bilingual participants. This study used a background questionnaire (one for Arabic-English bilinguals and one for the English-Arabic bilinguals) adapted from Schmid (2002), which contained two parts. The first part was used to gather demographic data from the participants, including age, gender, residence in different language environments, years of instruction and language qualifications. The second part collected self-reported data, such as contact times with speakers of each language, the frequency and type of usage of each language and the participants' attitudes towards the relevant languages and cultures. The second part gathered information about participants' L1 and L2 language use and contained questions that were measured on a five-point Likert scale to gauge how frequently participants used their L1 and L2 and with whom they used them; 100% refers to exclusive use of the language (Arabic or English) with particular people, 80% refers to frequent use, 50% refers sometimes using the language, 10% refers to seldom use and 0% to never using that language, as shown in example (7) (see Appendix B for more examples). Based on the bilinguals' answers to the background questionnaire, the participants were assigned to one of the two groups (more Arabic and more English).

(7) With whom did you talk Arabic after the time of your moving here? If you no longer speak Arabic with these people, please indicate until when you did so?

With your parents

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

With your parents until……………….

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

Language Proficiency and Aptitude

Proficiency Test

This study employed several excerpts from Arabic and English proficiency tests to obtain an indication of the subjects' proficiency levels in both Arabic and English. A computer-based version of the Test of English as a Foreign Language (TOEFL) and an Arabic proficiency test (APT) were used to assess the L2 language proficiency of all bilingual participants. The APT was provided by Umm Al-Qura University, the Arabic Language Institute, Makkah Al-Mukarramah, Saudi Arabia. The tests were limited to a grammar (structure) test and a listening comprehension test; both tests required knowledge of vocabulary and the ability to make grammatical judgments. The APT consisted of three sections containing 46 items. Section one of the test contained 15 multiple-choice questions and tested the participant's ability to understand the sound they heard. Section two contained 16 multiple-choice questions and tested the participant's listening comprehension ability. Section three contained 15 multiple-choice questions and tested the participants' knowledge of grammar (see Appendix A for example questions from each section). The test was scored by assigning one mark for each correct answer, using the answer guide provided with the test.

The online TOEFL test was used to test the participants' proficiency in English. The test contains three sections comprising 45 items. The first section comprises 15 items, involving written questions about brief audio recordings. The second section contains 10 items that asked participants to answer questions about a conversation and tested their listening comprehension ability. The third section contains 20 multiple-choice questions and tested the participants' knowledge of grammar (see Appendix A for example questions from all sections). The test was automatically scored, giving participants one point for each correct answer.

Language Aptitude Test

The language aptitude tests used in the current study were adapted from samples of the MLAT (Carroll & Span, 1959) and the Pimsleur Language Aptitude Battery (PLAB; Pimsleur, 1968) that are available online. The MLAT was designed to predict native English speakers' talent and success with L2 learning. Specifically, the test establishes the ability or individual talent of L2 learners. This study used the MLAT sample questions in Part III (Spelling Cues) and Part V (Paired Associates) and the PLAB sample questions in Part V

(Sound Discrimination). I then created questions analogous to the sample questions mentioned above to create three aptitude tests that are modelled on the MLAT's Spelling Cues and Paired Associates tests and the PLAB's Sound Discrimination test. When possible, I included the sample questions available online in the tests. A native speaker of the languages relevant for each test (English, Urdu and Cantonese) checked the test. All the three tests measure general language aptitude. I have selected these particular three tests because they involve knowledge about sounds, such as phonetic representations, or pattern recognition skills. In addition, a small pilot study confirmed the feasibility of the tests and ensured that none of the items were confusing for the participants.

The Spelling Cues test contains 25 questions (8 sample questions from the MLAT and 9 analogous questions that I created) that test a person's knowledge of English and his or her ability to associate sounds with symbols. Every question includes a group of words, and the target word placed at the top is spelled as it is pronounced or as the letters in the word are pronounced. The participants were told that to show that they recognised the disguised word, they had to recognise the word from the spelling (Carroll & Span, 1959). They were required to choose one of five words. For example, the term *grbj* must be linked with one of the related terms (see the MLAT sample question in example 8, where **boldface** indicates the correct response):

(8) *grbj*

A. car port
B. seize
C. boat
D. boast
E. waste

The Paired Associates test addresses the ability and memory aspects of learning a new language. Participants were presented with a list of 20 words in Urdu, a language unknown to them, and had two minutes to memorise the words. The original MLAT test presents vocabulary in Maya. I chose Urdu instead because I had access to native speakers of that language with a high proficiency in English who could check the tests. Once the participants began the test, they were not permitted to look at the vocabulary list again. They were then asked to translate the words that they had memorised, as shown in example 9.

(9) *kad-doo*

A. pumpkin

B. carrot
C. terrify
D. juice

The Sound Discrimination test requires students to distinguish pitch, orality and nasality in a language with which they are unfamiliar (PLAB; Pimsleur, 1968). I used Cantonese as the unfamiliar language because I had access to a female native speaker of that language who also had high English proficiency. The data recording took place in a soundproof room at Bangor University using a high-quality handheld Sony tape recorder. The participants first spent a few minutes learning the three Cantonese words *ngau* (Engl.: *cow*), *gau* (Engl.: *dog*) and *tau* (Engl.: *head*). To learn the three words, participants heard and saw each word three times with its English translation. Then, the participants' ability to recognise the words in a sentence was tested. They were exposed to 30 Cantonese sentences, each preceded by a beep, with each sentence containing one of these three words. The sentences were created, translated and produced by a native speaker of Cantonese who was highly fluent in English. Productions were recorded in a soundproof room. For each sentence, the participants were asked to circle the target word they heard in the sentence. For example, they heard the sentence in example 10.

(10) (audio recording)

佢	個	頭	好	大
keoi	go	tau	hou	daai

She has a big head.

(written answer choices)

A. Cow
B. Head

2.5.3 Overall Procedure

The data were collected over six months in KSA and the UK. The first process involved searching for and selecting the participants (A-E bilinguals and E-A bilinguals). As mentioned above, this was predominantly done using the snowball method.

The background questionnaire and consent form (see Appendices B and G) were emailed to the A-E bilinguals and E-A bilinguals along with an information sheet (see Appendix H), which included all relevant project information. Once consent to participate was obtained, a convenient and quiet place was selected as the location for the study. Ethical approval from Bangor University Research Ethics Committee, College of Arts and Humanities was received for the present study (see Appendix F).

The study comprised three sessions for bilingual participants, each on separate days up to one week apart. The first session began with the L2 proficiency assessment (either English or Arabic), followed by the aptitude test. The language proficiency assessment took approximately 30 to 40 minutes for each participant to complete, and the aptitude test took around 15 to 20 minutes to complete.

The second and third sessions involved three different production tasks and took about 10 to 15 minutes each. Arabic productions were recorded in one session, and English productions in the other. During each session, participants' speech was recorded with a high-quality handheld Sony tape recorder in a quiet environment. First, participants read a list of words and sentences, which contained different vowels (for the study reported in Chapter Three). Then participants read the question–answer dialogs analysed in the current study six times. Finally, participants narrated a short cartoon (for the study reported in Chapter Four), which allowed for the production of plosives. Monolingual participants only participated in the production tasks in their native language. Thus, their participation involved only one session.

2.5.4 Data Analysis

To analyse the pitch patterns on the wh-words, I first determined the beginning and end of each *wh*-word manually and then extracted the pitch contour of each *wh*-word automatically using Praat (Boersma & Weenink, 2016; see Figure 2.1 for an example of a waveform and spectrogram in Praat) and ProsodyPro (Xu, 2013). The automatically extracted pitch contours were hand-checked for possible tracking errors (such as pitch halving and pitch doubling) and hand-corrected if needed. Word duration (i.e. time) was normalised by taking 10 equally spaced measurements for each word. Pitch measurements were converted into the percentage of each individual's pitch range to facilitate the comparison of tokens. An individual's pitch range was determined manually from the complete question-answer dialog that each participant produced. The percentage of pitch range was calculated using the formula in (11), where *x* is the individual pitch value in Hertz, and *min* and *max* are the bottom (min) and top (max) of the particular participant's pitch range, respectively.

(11) ((x - min) * 100) / (max - min).

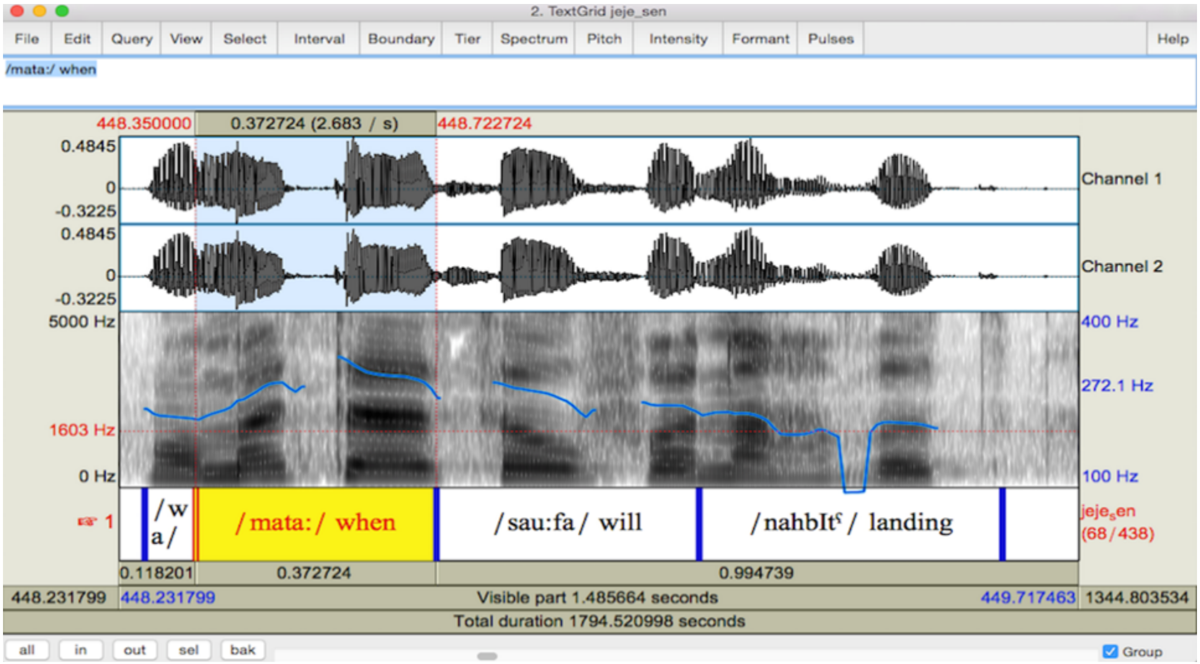


Figure 2.1. Example: waveform and spectrogram of monolingual native Arabic speaker

I used growth-curve analyses (GCA) (Mirman, 2014; Winter & Wieling, 2016) to analyse the pitch patterns on the *wh*-words. GCA is particularly suited to analysing patterns over time because it treats time in a statistically rigorous manner. As a type of mixed effects analysis, GCA can model both fixed and random effects. Here, I used GCA to analyse the pitch curves over time, in particular, during the production of the *wh*-words. I chose to use the entire *wh*-word as our analysis window for the GCA for several reasons: A smaller analysis window would run the risk of not capturing curvature. In particular, a GCA may model a small analysis window that only represents a portion of a curve as a linear model, i.e. as a straight line, rather than a curve. As a result, actual differences in curvature could potentially be obscured if only a portion of a curve is considered; for example, if only the stressed vowel of each wh-word was used as a window for analysis. Furthermore, the descriptions that I have of the pitch pattern for *wh*-words in Arabic do not provide us with any detailed temporal or segmental information that would allow us to systematically select a smaller analysis window.

The response variable for the GCAs shown below is the percent of pitch range (see above, formula in (11), for how this was calculated). All initial statistical models presented below contain the following fixed effects: up to three variables derived from the normalised time (see above for how word duration was normalised), the independent variable of interest (e.g. L1 vs L2) and the two-way interactions of the variable of interest with the different time variables. All initial models also included random intercepts for each participant as well as random slopes for the variable of interest and the time variables for each participant (see Winter & Wieling, 2016, for an explanation of fixed and random effects). Random intercepts and slopes were included to avoid Type I errors (see Barr et al., 2013). The particular statistical models presented below were determined through model comparisons (cf. Baayen, 2008) following Winter and Wieling (2016). I first established whether the full random effects structure was warranted. Random effects that did not contribute significantly to model fit or that caused the model to not converge were removed. I then determined whether the full fixed effects structure was warranted. Fixed effects that did not contribute significantly to model fit were removed from the model to derive the final statistical model.

Of particular relevance for the current chapter are the interactions of the variable of interest with the different time variables, as these provide information as to how the curves that are being compared differ in terms of the variable of interest. In particular, a significant interaction of the variable of interest with normalised time indicates that the levels of the variable of interest differ significantly in terms of steepness of the curves across the *wh*-word. A significant interaction of the variable of interest with squared normalised time means that the levels of the variable of interest differ significantly in terms of symmetric curvature (U-shaped or inverse U-shaped), i.e. in terms of how much the curve falls and rises or rises and falls around a central inflection point. A significant interaction of the variable of interest with cubed normalised time captures asymmetric curvature (sometimes in combination with a significant interaction of the variable of interest with normalised time), i.e. differences in terms of how much the curve rises and falls around an inflection point, where the rise and fall may be asymmetric. These differences are visualised in Figure 2.1, which is adapted from Mirman et al. (2008).

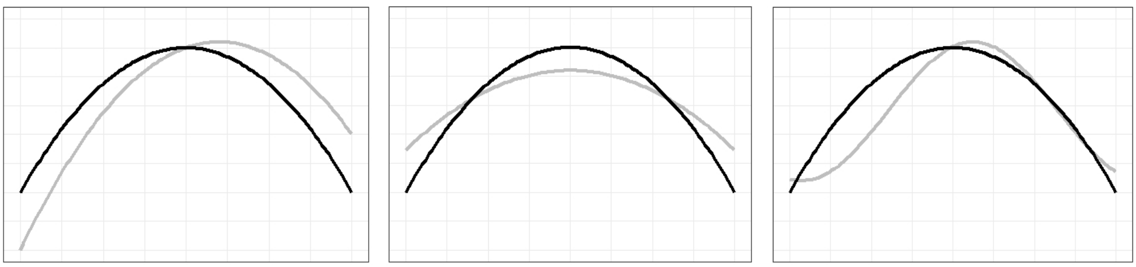


Figure 2.2. Visualisation of the differences that the interactions with the different time variables capture. Left: The interaction of the variable of interest with normalised time captures differences in slope. Middle: The interaction of the variable of interest with squared normalised time captures

differences in symmetric curvature. Right: The interaction of the variable of interest with cubed normalised time captures differences in terms of steepness around inflection points.

2.6 Results

2.6.1 Comparison of Monolingual Patterns

I first compare the pitch curves for the monolingual Arabic and English speakers to determine whether the pitch patterns of the *wh*-words differ between the two languages. Figure 2.3 shows the mean pitch curves for the Arabic and English *wh*-words. As speakers may have more time to realise a specific pitch pattern for longer words, Figure 2.3 shows the pitch curves for shorter, monosyllabic and longer disyllabic *wh*-words separately. The figure shows that there is a high target on the *wh*-word for disyllabic Arabic words. The pitch rises from the beginning of the word to the high target and then falls towards the end of the word. Monosyllabic Arabic *wh*-words show the same pattern, but it is less pronounced. English *wh*-words also show a slight rise in pitch, but there is no clear fall at the end of the word.

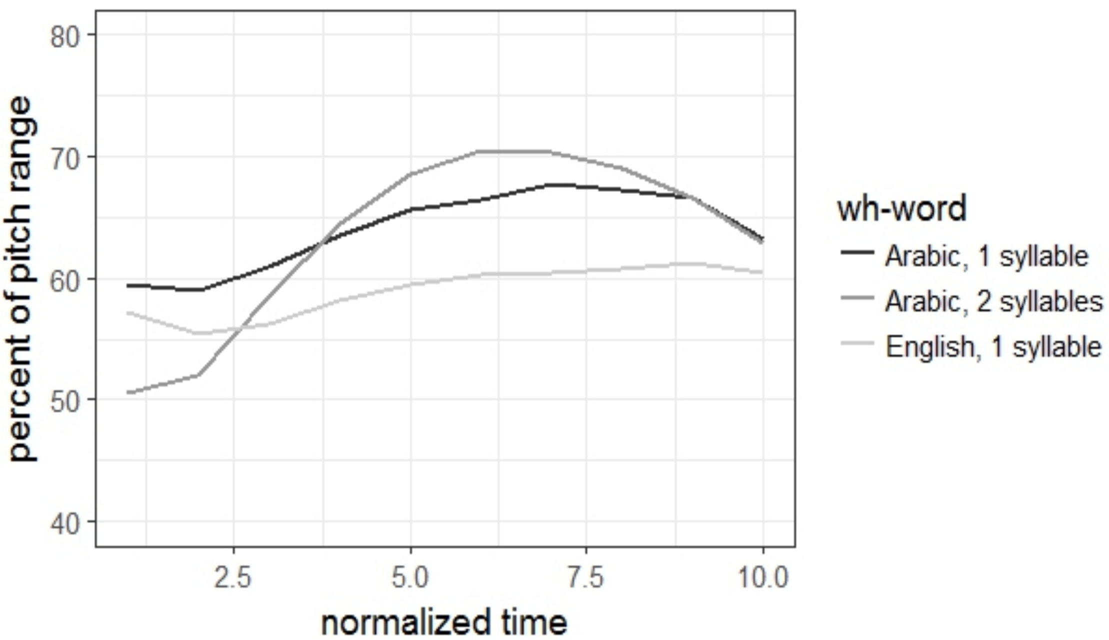


Figure 2.3. Mean pitch curves of Arabic and English monosyllabic and disyllabic wh-words

I performed a GCA to determine whether the pitch curves shown in Figure 2.3 differ reliably in terms of steepness or curvature. The response variable for the analysis was percent of pitch range. The fixed effects for the initial and final model were normalised time, squared normalised time, cubed normalised time, the type of *wh*-word (monosyllabic Arabic, disyllabic Arabic or monosyllabic English), and the two-way interactions of type of *wh*-word and the three time variables. Table 2.1 shows the results of the analysis. A significant effect of time was evident with a positive estimate, showing that, overall, F0 is higher at the end of the interval than at the beginning. Squared time was also significant with a negative estimate, revealing that the data overall correspond to an inverse U-shape. Cubed time was significant with a negative estimate, suggesting that the overall curve is steeper to the left of the inflection point than the right of the inflection point. Importantly, both the time by *wh*-word type and the squared time by *wh*-word type interactions were significant. The former effect suggests that the three *wh*-word types differ significantly in terms of steepness of the curve across the interval. The latter effect reveals that the three *wh*-word types differ reliably in terms of inverse U-shaped curvature.

Table 2.1. Results of the GCA Comparing the Three Types of Wh-words in Monolingual Arabic and English Speakers

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	65.18	3.47	31	18.76	< .001
Time	1.77	0.28	31	6.42	< .001
Squared time	-0.34	0.09	23	-3.7	< .001
Cubed time	-0.04	0.01	16160	-8.12	< .001
Wh-word type	0.41	1.98	19	0.21	= .837
Time * *wh*-word type	0.34	0.12	7146	2.93	=.003
Squared time * *wh*-word type	-0.18	0.03	1786	-5.42	< .001
Cubed time * *wh*-word type	0.01	0.01	16160	1.51	= .131

I performed post-hoc GCA to explore which of the *wh*-word types differ significantly from one another in terms of steepness of the curve and in terms of inverse U-shaped curvature. The focus of the post-hoc analysis is performed to determine whether the curves for the monosyllabic and disyllabic Arabic *wh*-words differ significantly from the curves for the monosyllabic English *wh*-words. I first compare monosyllabic Arabic words with monosyllabic English words. The initial and final model for this analysis included the percent of pitch range as response variable and normalised time, squared normalised time, *wh*-word type and the two-way interactions of *wh*-word type with the two time variables as fixed effects. Table 2.2 presents the results. The relevant results are the two interactions. The time by *wh*-word type interaction was not significant, suggesting that the monosyllabic Arabic words do not differ reliably from the monosyllabic English words in terms of steepness of the slope across the interval. That is, the lines of best fit for both curves do not differ in terms of steepness. The squared time by *wh*-word type interaction was significant, showing that the monosyllabic Arabic words differ reliably from the monosyllabic English words in terms of curvature. In particular, the curve for the monosyllabic Arabic words is significantly more inverse U-shaped, i.e. is more strongly curved, than the curve for the monosyllabic English words.

Table 2.2. Post-hoc GCA Comparing Monosyllabic Arabic and Monosyllabic English Wh-words

Fixed effects	**Estimate**	**std. error**	**Df**	***t*-value**	***p*-value**
Intercept	60.59	3.52	26.6	17.21	< .001
Time	0.63	0.17	14.79	3.65	< .01
Squared time	-0.07	0.03	12.02	-2.32	< .05
Wh-word type	-1.75	1.68	27.17	-1.04	= .313
Time * *wh*-word type	-0.06	0.09	22.35	-0.695	= .486
Squared time * *wh*-word type	0.06	0.02	50.39	2.59	= .018

Next, I compare disyllabic Arabic words with monosyllabic English words. The final model for this analysis also included the percent of pitch range as the response variable and normalised time, squared normalised time, *wh*-word type and the two-way interactions of *wh*-word type with the two time variables as the fixed effects. Table 2.3 shows the results. The relevant results are the two interactions. The time by *wh*-word type interaction was significant, showing that the disyllabic Arabic words differ reliably from the monosyllabic English words in terms of steepness of the slope across the interval. That is, the lines of best fit for the curve for disyllabic Arabic words are steeper than those for the monosyllabic English words. The squared time by *wh*-word type interaction was also significant, showing that the disyllabic Arabic words differ reliably from the monosyllabic English words in terms of symmetric inverse U-shaped curvature. The curve for the monosyllabic Arabic words is significantly more inverse U-shaped, i.e. more strongly curved, than is the curve for the monosyllabic English words.

Table 2.3. Post-hoc GCA Comparing Disyllabic Arabic and Monosyllabic English Wh-words

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	64.26	2.84	27	22.63	< .001
Time	1.13	0.22	26.92	5.24	< .001
Squared time	-0.34	0.04	26.58	-7.98	< .001
Wh-word type	-4.39	2.83	27	-1.55	= .132
Time * *wh*-word type	-0.54	0.22	26.92	-2.512	= .030
Squared time * *wh*-word type	0.3	0.04	26.59	6.91	< .001

Overall, the results confirm the observations of the literature of a steep rise to a high target in Arabic *wh*-words, but no such high target in English *wh*-words. The high target is realised more clearly in the disyllabic Arabic words, which differ significantly from monolingual English wh-words both in terms of a steeper rise across the interval and a stronger inverse U-shape. Importantly, the shorter monosyllabic Arabic *wh*-words, where time is limited to realise a high target, also differ from monosyllabic English *wh*-words in terms of a stronger inverse U-shape. This finding suggests that they are also more likely to have a high target compared to monosyllabic English wh-words. To summarise and reiterate, the significant squared time by *wh*-word type interactions above show that the Arabic *wh*-words have a significantly steeper rise-fall pattern (inverse U-shape) than do the English *wh*-words. Having established a reliable difference in pitch curve patterns between Arabic and English *wh*-words in monolingual speakers of the two languages, I move on to examining bilingual speakers.

2.6.2 Profile of Bilingual Participants

I start the analyses of bilingual speakers by providing a profile of the bilingual participants. In particular, I provide information about the bilingual participants' L2 proficiency, their phonetic aptitude and their L1 and L2 language use. Table 2.4 presents a summary of the information, indicating that both groups are highly proficient, with a mean TOEFL score of 42.1 out of a possible 45 for the A-E bilinguals and a mean APT score of 43.7 out of 46 for the E-A bilinguals. The two-sample t-tests reported in Table 2.4 show that the two bilingual groups did not differ in their performance in any of the phonetic aptitude tests. For the purposes of further analyses, I divided A-E and E-A bilinguals into two aptitude groups based on their aptitude test results. Nine A-E bilinguals and six E-A bilinguals were grouped as low aptitude, with scores between 47 to 54 out of a possible 75. The remaining six A-E bilinguals and eight E-A bilinguals were considered to be high aptitude, with scores between 57 and 71. For the purposes of the following analyses, I also divided A-E and E-A bilinguals into two language use categories based on the comprehensive language background questionnaire. Eight A-E bilinguals and four E-A bilinguals were categorised as more Arabic, meaning they used Arabic as frequently as English or more frequently than English. The remaining seven A-E bilinguals and 10 E-A bilinguals were classed as more English, meaning they used English more than Arabic. Thus, overall, the A-E bilinguals tend to use more Arabic than do the E-A bilinguals.

Table 2.4. Proficiency, Aptitude and Language Use of A-E and E-A Bilinguals

Measure	**A-E bilinguals**	**E-A bilinguals**	***t*-test *p*-value**
Proficiency	TOEFL: mean = 42.1 (SD = 1.9; range = 39-45)	APT: mean = 43.7 (SD = 1.0; range = 42-46)	NA
Aptitude spelling cues	mean = 17.5 (SD = 4.2; range = 11-25)	mean = 19.5 (SD = 3.3; range = 14-24)	$t = -1.4494$ $p = 0.16$
Aptitude paired associates	mean = 14.6 (SD = 2.0; range = 12-18)	mean = 15.7 (SD = 2.2; range = 12-19)	$t = -1.4155$ $p = 0.17$
Aptitude sound discrimination	mean = 23.5 (SD = 2.3; range = 20-28)	mean = 23.4 (SD = 2.7; range = 20-28)	$t = 0.1883$ $p = 0.85$
Aptitude total	mean = 55.6 (SD = 7.8; range = 45-71)	mean = 58.6 (SD = 7.5; range = 49-69)	$t = -1.0437$ $p = 0.31$
Language use	More Arabic = 8 More English = 7	More Arabic = 4 More English = 10	NA

Note: TOEFL = Test of English as a Foreign Language; APT = Arabic Proficiency test.

2.6.3 Bilinguals: Acquisition

Comparison of E-A bilinguals with Arabic Monolinguals in Arabic

First, I investigate how well the A-E and E-A bilinguals acquire the target prosodic pattern on the *wh*-word in their L2 and determine whether the acquisition of this pattern is influenced by phonetic aptitude and/or language use. Based on the results above that compare monolingual English and monolingual Arabic speakers, I focus on the differences in the symmetric inverse U-shaped curvatures between the groups. I start by comparing the pitch curves of Arabic monolinguals and E-A bilinguals in Arabic to determine whether the pitch contours on the *wh*-words differ across the two groups. In particular, if the E-A bilinguals showed a reliably less symmetric inverse U-shaped curvature than that of the Arabic monolinguals, I would consider this evidence of interference from the L1 English. This would be reflected in a significant interaction between squared time and group (E-A bilinguals vs Arabic monolinguals). Figure 2.4 shows the mean pitch curves for Arabic monolinguals and E-A bilinguals of *wh*-words in Arabic. The figure shows that there is a high target on the *wh*-words for Arabic monolinguals: the pitch rises from the beginning of the word to a high target and then falls towards the end of the word. E-A bilinguals show a similar, but somewhat less curved pattern.

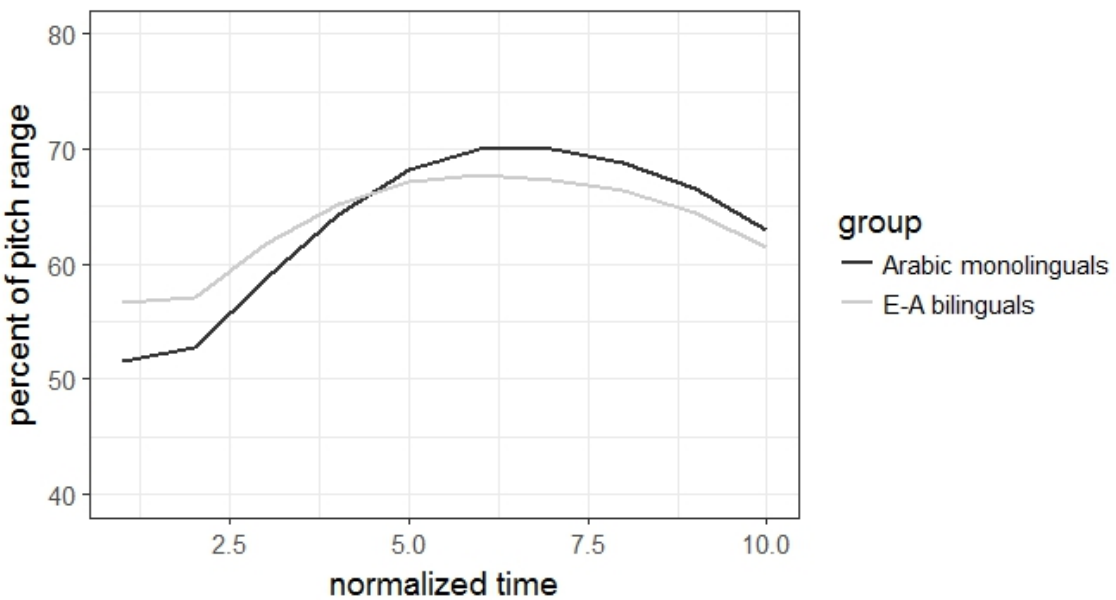


Figure 2.4. Mean pitch contours of wh-words in Arabic produced by Arabic monolinguals and E-A bilinguals.

I performed a GCA with percent of pitch range as the response variable, and group, the three time variables and all two-way interactions of group with the time variables as the fixed effects to determine whether the pitch contours of *wh*-words in Arabic produced by Arabic monolinguals and E-A bilinguals differed reliably in terms of steepness or curvature. Table 2.5 shows the results of the complete analysis. For this and all following analyses, I only report relevant interactions with the time variables in the text because these reflect differences in steepness or curvature across the compared groups. All other results can be found in the accompanying tables, but will not be mentioned in the text. The results from the analysis in Table 2.5 reveal that the time by group interaction was significant, suggesting that the groups differed in terms of steepness across the interval with a steeper rise across the interval for Arabic monolinguals compared to E-A bilinguals. The squared time by group interaction was not significant, suggesting that the Arabic monolinguals showed almot similar inverse U-shaped curvature as the E-A bilinguals. Finally, the cubed time by group interaction was significant, suggesting that the groups differ in terms of steepness around the inflection point. Overall, the pattern of results suggests that the E-A bilinguals are approaching the native pattern.

Table 2.5. Results of the GCA Comparing Wh-words in Arabic across Arabic Monolinguals and E-A Bilinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	**t-value**	**p-value**
Intercept	68.00	2.54	27	26.77	< .001
Time	1.80	0.22	36	8.17	< .001
Squared Time	-5.24	2.70	27	-11.11	< .001
Cubed Time	-0.03	0.005	1526	-7.41	< .001
Group	-0.49	2.52	27	-0.19	= .846
Time * group	-0.60	0.21	36	-2.73	= .010
Squared Time * group	0.08	0.04	27	1.89	= .069
Cubed Time * group	0.01	0.005	1526	2.37	= .018

Since the E-A bilinguals differed reliably from the Arabic monolinguals in their productions of the *wh*-words, I explored whether phonetic aptitude or language use modulated the extent to which the E-A bilinguals approximated the native Arabic pitch pattern. For this and similar following analyses, I divided the participants into those with low and high phonetic aptitude and into those who use more Arabic and more English in their everyday lives (see the section 2.7.2 on the profile of bilingual participants and Table 2.5 above). I explored the possible effects of phonetic aptitude and language use separately because none of the participants were categorised as both more Arabic and low aptitude. Here, I first examined whether language use has an effect on the pitch contour of *wh*-words in Arabic as produced by high L2 use (more Arabic) and low L2 use (more English) E-A bilinguals. Figure 2.5 indicates that the high L2 use participants show more inverse U-shaped curvature and a peakier pattern than do the low L2 use participants.

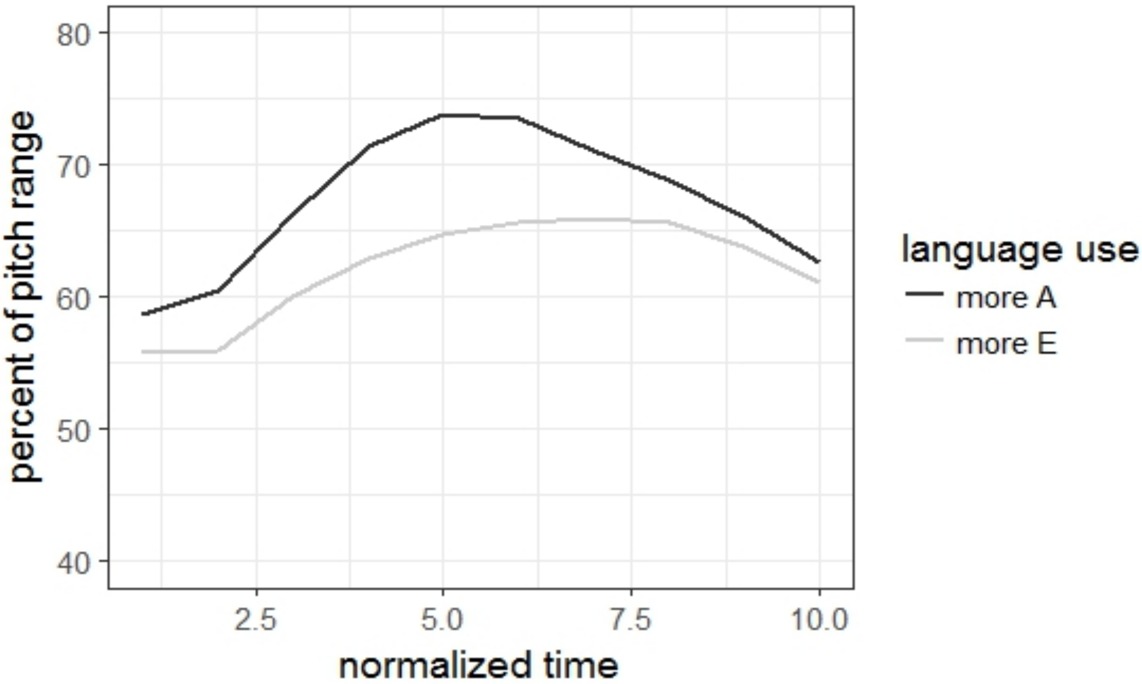


Figure 2.5. Mean pitch contours of wh-words in Arabic according to language use of E-A bilinguals.

I performed a GCA with percent of pitch range as the response variable and language use, the three time variables and all two-way interactions of language use and the time variables as the fixed effects for E-A bilinguals to analyse whether the pitch contours differed reliably by language use in terms of curvature or steepness. The statistical results presented in Table 2.6 show that the time by language use interaction was not significant, suggesting a slightly steeper line of best fit for the low L2 use (more English) participants compared to the high L2 use (more Arabic) participants. Both the squared and cubed time by language use interactions were significant, which suggests that the participants with high L2 use showed stronger symmetric inverse U-shaped curvature, i.e. a more native-like pattern, than did the participants with low L2 use. In addition, the reliably peakier pattern produced by the high L2 use participants compared to the low L2 use participants suggests that the high L2 use participants might be overshooting the target.

Table 2.6. Results of the GCA Comparing Wh-words in Arabic Produced by E-A bilinguals in the High and Low L2 Use Groups

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	67.3	2.51	12	26.80	< .001
Time	1.15	0.18	23	6.27	< .001
Squared Time	-0.42	0.05	12	-8.24	< .001
Cubed Time	-0.03	0.00	6936	-3.93	< .001
Language use	-2.83	2.46	12	-1.15	= .272
Time * language use	0.36	1.80	24	2.00	= .056
Squared Time*language use	0.01	0.05	12	2.41	= .032
Cubed Time * language use	-0.01	0.00	6926	-2.39	= .016

I then examined whether aptitude has an effect on the pitch contour of *wh*-words in Arabic as produced by high and low aptitude E-A bilinguals. Figure 2.6 shows the mean pitch curves of *wh*-words in Arabic produced by E-A bilinguals with high and low phonetic aptitudes. The curves at the beginning are nearly identical between the two groups. However, the high aptitude E-A bilinguals showed a clear rise and fall pattern in Arabic *wh*-words, while the low aptitude group showed a clear rise with a later peak and a less clear fall at the end.

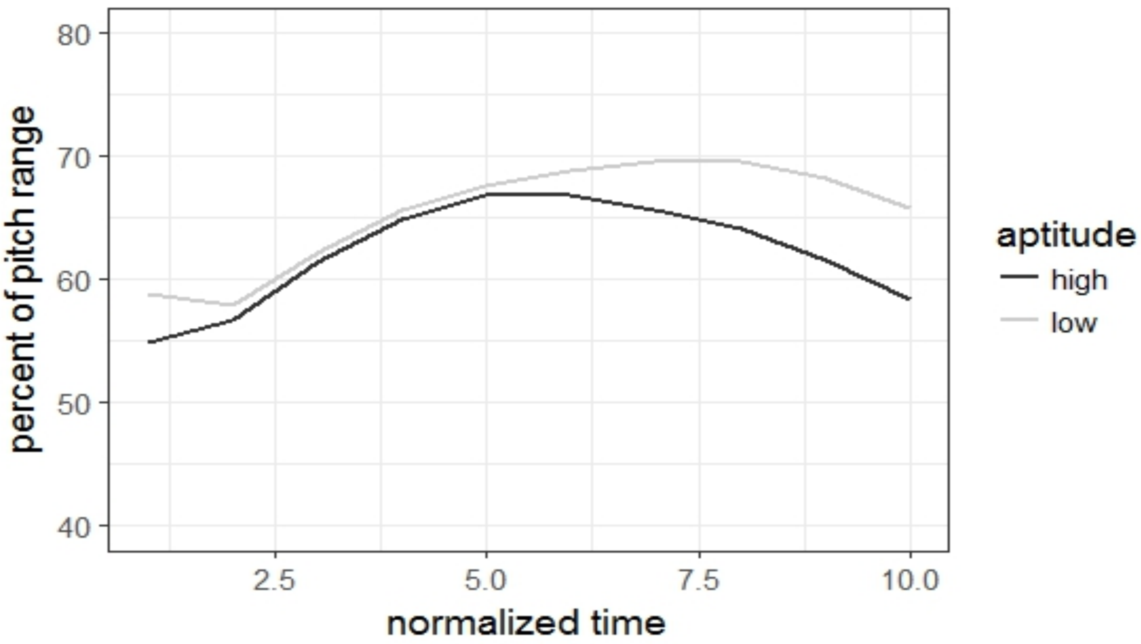


Figure 2.6. Mean pitch contours of wh-words in Arabic for low and high aptitude E-A bilinguals

I performed a GCA with percent of pitch range as the response variable and phonetic aptitude, the three time variables and all two-way interactions of aptitude and the time variables as the fixed effects to determine whether aptitude influenced the pitch pattern for *wh*-words in English produced by E-A bilinguals. The statistical results summarised in Table 2.7 show that the time by aptitude and cubed time by aptitude interactions were significant, suggesting that a difference exists between the aptitude groups in terms of the steepness of the curve across the interval and the steepness around inflection points. The squared time by aptitude interaction was not significant, revealing that no difference exists between the high and low aptitude groups in terms of symmetric inverse U-shaped curvature. Overall, aptitude clearly modulated the produced pitch patterns, but I found no evidence that one of the patterns could be considered more native-like than the other or that one pattern was characterised more by L1 interference than the other. In particular, both patterns showed similar symmetric inverse U-shaped curvature, but seemed to differ in the timing of the peak.

Table 2.7. Results of GCA Comparing Wh-words in Arabic Produced by High and Low Aptitude E-A Bilinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	67.45	2.61	12	25.76	< .001
Time	0.16	1.66	28	6.84	< .001

Squared Time	-0.42	0.05	12	-7.62	< .001
Cubed Time	-0.02	0.00	6936	-3.93	< .001
Aptitude	1.15	2.61	12	0.44	= .665
Time * Aptitude	0.58	0.16	28	3.53	< .01
Squared Time * Aptitude	0.09	0.05	12	1.68	= .119
Cubed Time * Aptitude	-0.02	0.00	6936	-3.07	< .01

Comparison of A-E bilinguals with English Monolinguals in English

Next, I compare the pitch curves of English monolinguals and A-E bilinguals in English to determine whether pitch contours on *wh*-words differ between the two groups. Figure 2.7 shows that English monolinguals pronounced *wh*-words with a slight rise in pitch, while A-E bilinguals pronounced them with a more marked rise in pitch. Neither group exhibited a clear fall at the end of *wh*-words.

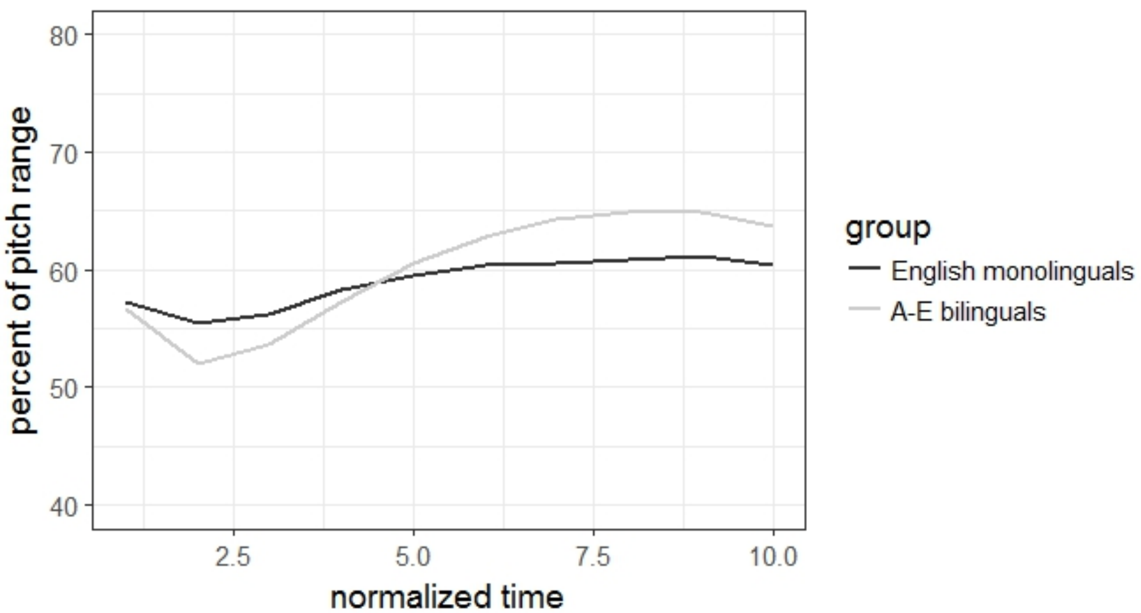


Figure 2.7. Mean pitch contours of wh-words in English produced by A-E bilinguals and English monolinguals.

I conducted a GCA with percent of pitch range as the response variable, and group, the three time variables and all two-way interactions between group and the time variables as the fixed effects to examine whether the pitch contours differed in terms of curvature or steepness between the two groups. Model comparisons showed that squared time and its interaction with group could be removed from the model. Both fixed effects were therefore not included in the final model reported in Table 2.8. Both the time by group and cubed time by group interactions were significant, suggesting that the two groups exhibit differences in the steepness of the pitch curve across the interval and the steepness around inflection points. These differences between the two groups suggest that the A-E bilinguals are approaching the native pattern, but that the differences cannot be attributed to L1 transfer or interference. In particular, A-E bilinguals do not differ from English monolinguals in terms of symmetric inverse U-shaped curvature, but they show a greater pitch excursion mostly characterised by a steeper rise in pitch throughout the interval than do English monolinguals.

Table 2.8. Results of the GCA Comparing Wh-words in English Produced by A-E Bilinguals and English Monolinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	59.93	2.40	27	25.26	< .001
Time	2.02	0.18	49	12.08	< .001
Cubed Time	0.07	0.006	1547	-11.80	< .001
Group	-0.001	2.37	27	-0.00	= .999
Time * Group	-0.90	0.18	49	-5.37	< .001
Cubed Time * Group	0.03	0.00	1547	5.70	< .001

Since the A-E bilinguals differed reliably from English monolinguals in their productions of the *wh*-words, I now explore whether phonetic aptitude or language use modulate the extent to which A-E bilinguals approximate the native English pitch pattern. Figure 2.8 shows the pitch curves for A-E bilinguals with high (more English) and low (more Arabic) L2 language use. Both patterns show a clear rise in pitch throughout the interval. However, this rise seems to be steeper for the low L2 (more Arabic) language use bilinguals than the high L2 (more English) bilinguals.

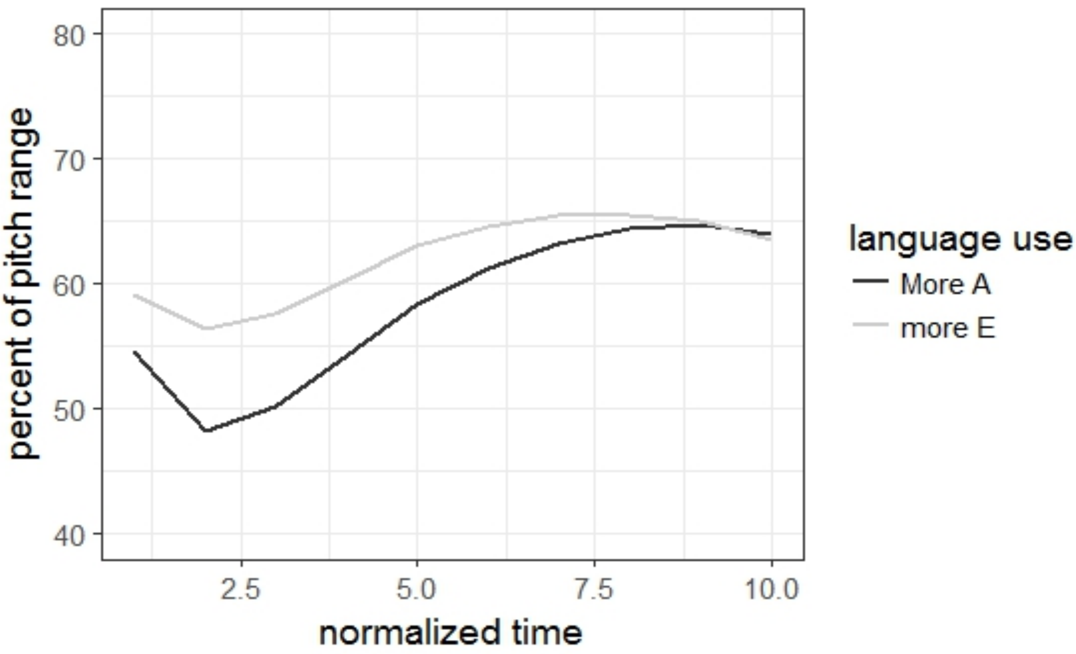


Figure 2.8. Mean pitch contours of wh-words in English produced by A-E bilinguals with low and high L2 language use

I conducted a GCA with percent of pitch range as the response variable, and language use, the three time variables and all two-way interactions of language use and time variables as the fixed effects to determine whether the pitch contours on *wh*-words produced by A-E bilinguals with high and low L2 use differed reliably in terms of steepness or curvature. Again, the model comparisons showed that squared time and its interaction with language use could be removed from the model. These fixed effects were therefore not included in the final model reported in Table 2.9. Both the time by language use and cubed time by language use interactions were significant, suggesting that a reliable difference exists between language use groups in terms of the steepness of the curve across the interval and steepness around inflection points. A-E bilinguals with low L2 use showed a reliably steeper rise throughout the interval than those with high L2 use. Considering that the English monolingual pattern has only a very slight rise in pitch, it seems that the high L2 use participants are approximating the native pattern more closely than are the low L2 use participants. Notably, neither group are approximating the Arabic symmetric rise-fall pattern.

Table 2.9. Results of the GCA Comparing Wh-words in English Produced by A-E bilinguals with High and Low L2 Use

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	59.96	3.36	13	17.85	< .001
Time	2.93	2.68	22	10.92	< .001
Cubed Time	-0.10	0.00	7612	-11.49	< .001
Language use	2.65	3.36	13	0.78	= .444
Time * Language use	-0.77	0.26	22	-2.89	< .01
Cubed Time * Language use	0.02	0.00	7612	2.67	< .01

Next, I consider phonetic aptitude. Figure 2.9 shows the pitch curves for A-E bilinguals with high and low phonetic aptitude. Both patterns show a clear rise in pitch throughout the interval.

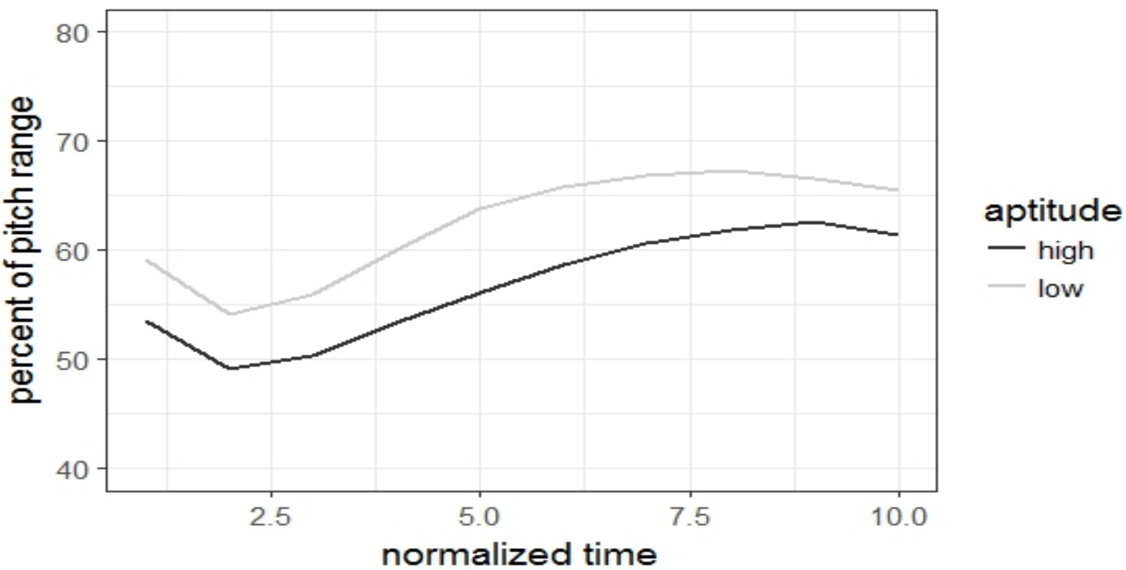


Figure 2.9. Pitch contours of wh-words in English produced by high and low aptitude A-E bilinguals

I conducted a GCA with percent of pitch range as the response variable and phonetic aptitude, the three time variables and all two-way interactions between aptitude and time variables as the fixed effects to determine whether aptitude influenced the curvature of *wh*-words in the English produced by A-E bilinguals. Model comparisons showed that squared time and its interaction with aptitude could be removed from the model and are thus not part of the final model. The results in Table 2.10 showed neither a significant interaction between time and aptitude nor between cubed time and aptitude. Thus, aptitude did not affect A-E bilinguals' pitch patterns.

Table 2.10. Results of GCA Comparing Wh-words in English Produced by High and Low Aptitude A-E Bilinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	60.66	0.27	7456	224.97	< .001
Time	2.94	0.30	23	9.74	< .001
Cubed Time	-0.10	1.01	7626	-10.03	< .001
Aptitude	3.33	0.27	7462	12.36	< .001
Time * aptitude	-0.01	0.30	23	-0.04	= .969
Cubed Time * aptitude	-0.00	0.01	7626	-0.28	= .779

Overall, the acquisition results suggest that both E-A and A-E bilinguals are approximating the native English and native Arabic patterns. Language use, but not – or at least to a much lesser extent – phonetic aptitude, modulates how closely bilingual participants approximate the native patterns.

2.6.4 Bilinguals: L1 Attrition

Comparison of A-E bilinguals with Arabic Monolinguals in Arabic

Next, I search for evidence of L1 attrition in the A-E bilinguals' Arabic compared to monolingual Arabic speakers. I focus again on differences in the symmetric inverse U-shaped curvature between the groups. In particular, if the A-E bilinguals showed a reliably less symmetric inverse U-shaped curvature than the Arabic monolinguals, I would take this as evidence of L1 attrition caused by the influence of L2 English. This would be reflected in a significant interaction between squared time and group (A-E bilinguals vs Arabic monolinguals). Figure 2.10 shows the mean pitch contours of the Arabic *wh*-words produced by A-E bilinguals and Arabic monolinguals. The curves for both groups are rather similar, both displaying a rise in pitch, a high target and a following fall on *wh*-words.

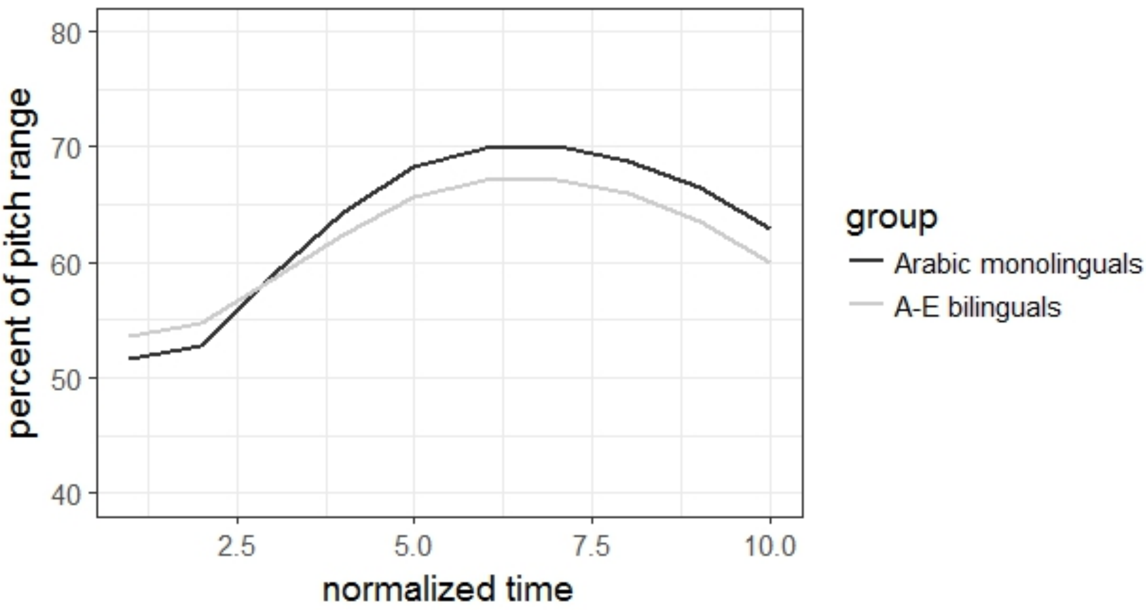


Figure 2.10. Mean pitch contours of wh-words for Arabic monolinguals and A-E bilinguals

A GCA was performed to determine whether the pitch contours shown in Figure 2.10 differ reliably in terms of steepness or curvature. Percent of pitch range served as the response variable for the analysis. The fixed effects for the initial and final models were normalised time, squared normalised time, cubed normalised time, group (Arabic monolinguals and A-E bilinguals) and the two-way interactions between group and the three time variables.

Table 2.11 shows that none of the time terms interacted significantly with group, suggesting that the pitch patterns of the A-E bilinguals and the Arabic monolinguals do not differ reliably in terms of steepness or curvature. I thus find no evidence of L1 attrition in terms of the pitch patterns on *wh*-words in the A-E bilinguals.

Table 2.11. Results of the GCA Comparing Wh-words for Arabic Monolinguals and A-E Bilinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	67.06	2.79	28	24.03	< .001
Time	2.10	0.23	38	9.33	< .001
Squared Time	-0.54	0.05	28	-9.98	< .001
Cubed Time	-0.05	0.01	16180	-9.08	< .001
Group	-1.44	2.79	28	-0.52	= .611
Time * Group	-0.26	0.22	38	-1.15	= .285

Squared Time * Group	0.07	0.05	28	1.27	= .215
Cubed Time * Group	-0.00	0.01	16180	-0.22	= .829

Comparison of E-A Bilinguals with English Monolinguals in English

Next, I consider whether there is any evidence of L1 attrition in the E-A bilinguals' English when compared to monolingual English speakers. Again, I focus on differences in the symmetric inverse U-shaped curvature between the groups. In particular, if the E-A bilinguals showed a reliably more symmetric inverse U-shaped curvature than the English monolinguals, I would take this as evidence of L1 attrition due to the influence of L2 Arabic. Figure 2.11, which illustrates the pitch contours for both groups, shows a slight rise and fall pattern on the *wh*-words for the E-A bilinguals, but a slight rise with no fall for the English monolinguals.

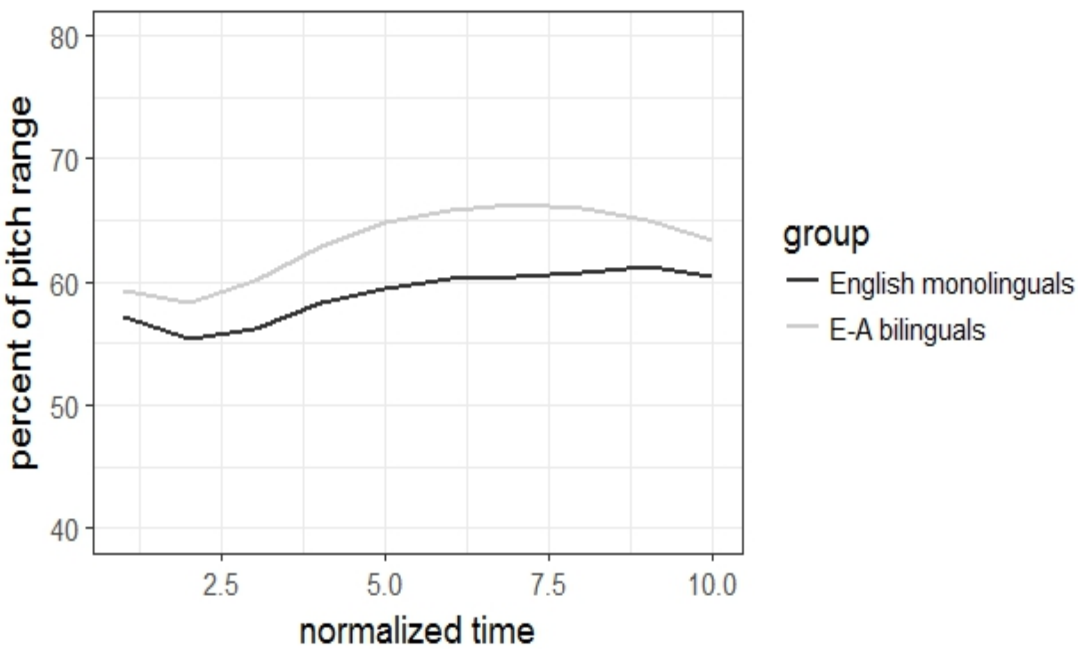


Figure 2.11. Pitch contours of English wh-words for English monolinguals and E-A bilinguals

To determine whether the pitch contours differed reliably in terms of steepness or curvature for the E-A bilinguals compared to the English monolinguals, I performed a GCA analogous to that for A-E bilinguals and Arabic monolinguals above. Table 2.12 presents the results. The time by group and the cubed time by group interactions were not significant, suggesting no group difference in terms of steepness of the curve across the interval and steepness around inflection points. However, the squared time by group interaction was significant, indicating that the English monolinguals differ reliably from the E-A bilinguals in terms of the symmetric inverse U-shaped curvature. In particular, the E-A bilinguals show a reliably stronger curvature compared to the English monolinguals, which suggests that the E-A bilinguals are more likely to have a high target on the *wh*-word, thus providing evidence for L1 attrition in the E-A bilinguals due to the influence of Arabic.

Table 2.12. Results of the GCA Comparing Wh-words for English Monolinguals and E-A Bilinguals

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	62.18	2.25	26	27.5	< .001
Time	1.30	0.13	78	9.99	< .001
Squared Time	-0.12	0.02	25	-5.08	< .001
Cubed Time	-0.04	0.005	14900	-7.30	< .001
Group	2.30	2.25	26	1.02	= .316
Time * Group	0.17	0.12	79	1.34	= .184
Squared Time * Group	-0.07	0. 02	25	-2.9	< .001
Cubed Time * Group	-0.006	0.005	14900	-1.10	= .268

Since the E-A bilinguals showed evidence of L1 prosodic attrition, I explored whether the strength of attrition was modulated by phonetic aptitude or language use. Figure 2.12 shows the mean pitch contours of *wh*-words in English produced by the more Arabic and more English groups of E-A bilinguals. The figure shows that E-A bilinguals with high L2 use (more Arabic) have a slightly steeper rise in their *wh*-question words than do those with low L2 use (more English).

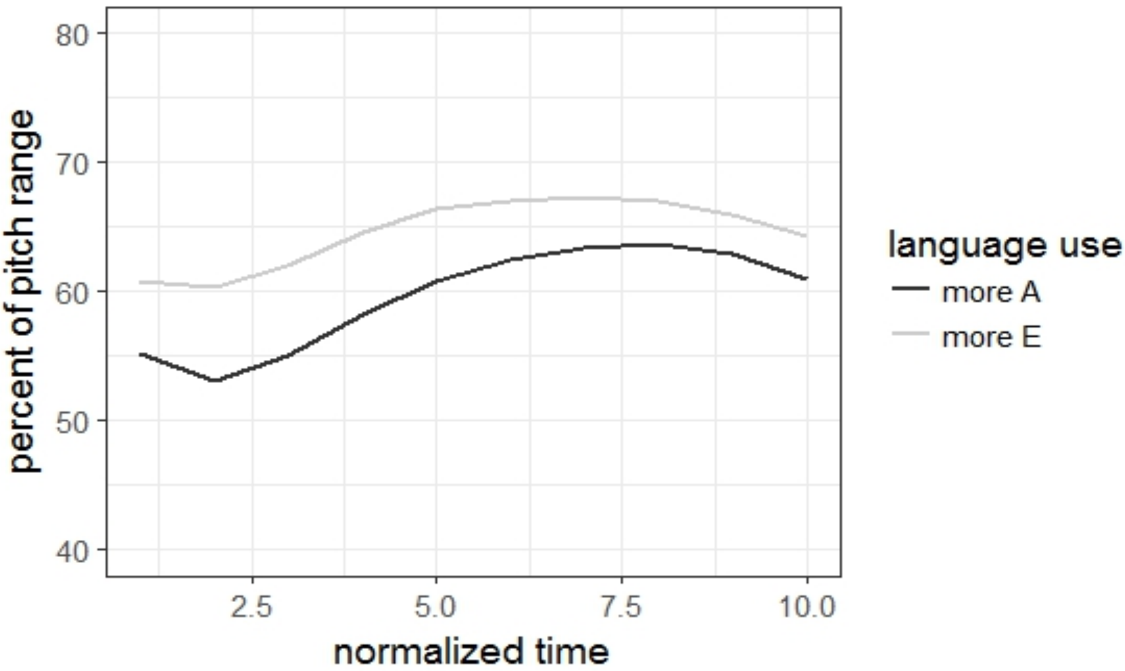


Figure 2.12. Mean pitch contours of wh-words in English produced by E-A bilinguals with low (more English) and high (more Arabic) L2 use

I performed a GCA for language use with percent of pitch range as the response variable and language use, the three time variables and the interactions of aptitude with the time variables as the fixed effects to determine whether E-A bilinguals' language use influenced the pitch curvature on *wh*-words in the L1 English. Table 2.13 shows that only the time by language use interaction was significant, suggesting a steeper rise across the interval for the high L2 use group compared to the low L2 use groups. Notably, the squared time by language use interaction was not significant, revealing that no difference exists between the low and high L2 use groups in terms of the symmetric inverse U-shaped curvature. Thus, I find no evidence that language use influences the strength of attrition.

Table 2.13. Results of the GCA Comparing Wh-words in English for E-A Bilinguals with Low and High L2 Use

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	64.65	3.02	12	21.40	< .001
Time	1.50	0.19	39	7.58	< .001
Squared Time	-0.19	0.03	12	-5.44	< .001
Cubed Time	-0.04	0.008	7046	-5.53	< .001
Language Use	2.27	2.99	12	0.75	= .462
Time * language use	-0.43	0.19	40	-2.23	=.030

Squared Time*language use	-0.01	0.03	12	-0.42	= .681
Cubed Time *language use	0.01	0.008	7046	1.69	= .092

Figure 2.13 shows the mean pitch contours of *wh*-words in English produced by the low and high aptitude E-A bilinguals. Both groups showed similar curves, but the low aptitude group produced the *wh*-words somewhat higher in their pitch range than the high aptitude group.

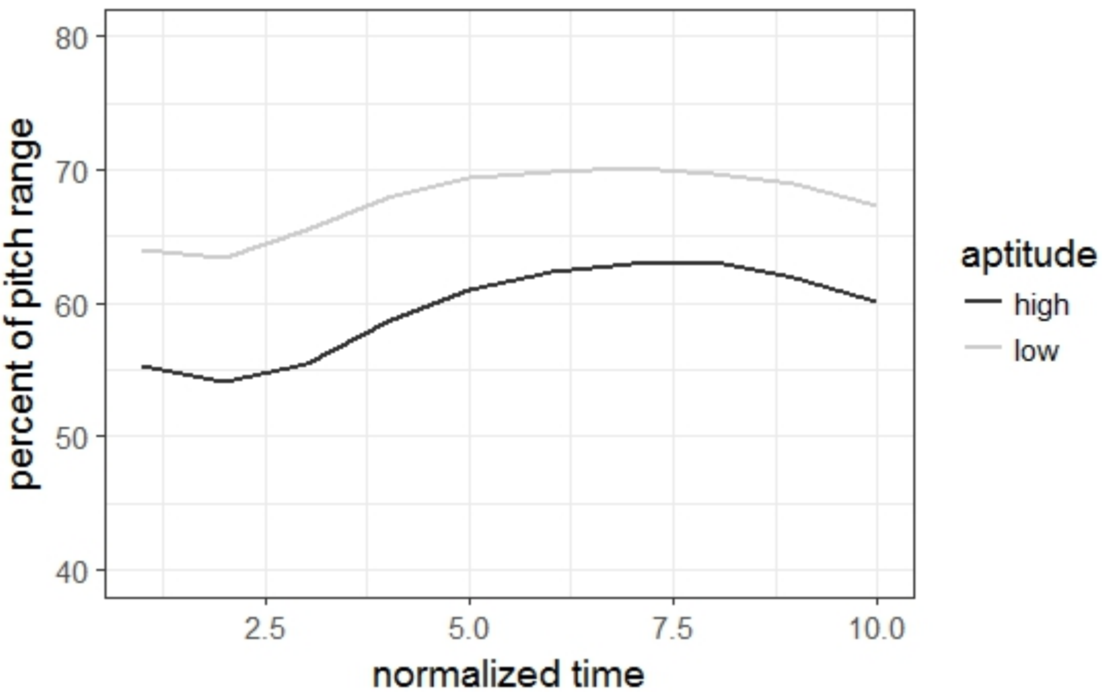


Figure 2.13. Mean pitch contours of wh-words in English produced by high and low aptitude E-A bilinguals.

I conducted a GCA for E-A bilinguals analogous to that performed for language use above to determine whether phonetic aptitude influenced the curvature of *wh*-words in the English produced by E-A bilinguals. The results in Table 2.14 show that the time by aptitude and cubed time by aptitude interactions were not significant, indicating no difference between the steepness of the curve across the interval and steepness around inflection points. Importantly, the squared time by aptitude interaction was not significant, which suggests that no difference exists between the high and low aptitude groups in terms of the symmetric inverse U-shaped curvature. Thus, I find no evidence that phonetic aptitude influences the strength of attrition.

Table 2.14. Results of the GCA Comparing Wh-words in English Produced by E-A Bilinguals with Low and High Phonetic Aptitude

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	64.83	2.86	12	22.60	< .001
Time	1.47	0.199	36	7.39	< .001
Squared Time	-0.19	0.04	12	-5.40	< .001
Cubed Time	-0.05	0.009	7046	-5.53	< .001
Aptitude	4.09	2.88	12	1.42	= .180
Time * Aptitude	-0.38	0.20	36	-1.91	= .063
Squared Time * Aptitude	-0.005	0.037	12	-0.16	= .873
Cubed Time * Aptitude	0.015	0.008	7046	1.70	= .088

Overall, the results for bilingual participants in their L1 reveal an asymmetry in terms of L1 prosodic attrition, with A-E bilinguals showing no evidence of attrition in their native Arabic but E-A bilinguals showing attrition in their native English. No evidence was found to indicate that the strength of attrition is influenced by phonetic aptitude or language use.

2.7 Discussion

This study investigated prosodic acquisition and L1 prosodic attrition in the production of *wh*-question words in late E-A and A-E bilinguals, that is, in bilinguals who speak two typologically different languages. In addition, this study explored whether bilinguals' phonetic aptitude and/or language use influenced how target-like their prosody is in the L1 and the L2. In this study, the bilinguals were directly compared to the monolinguals. For example, A-E and E-A bilinguals were compared with English monolinguals. The first comparison concerns L2 acquisition as bilinguals with English as an L2 are being compared directly with English monolingual speakers. The second comparison concerns L1 attrition as bilinguals with English as L1 are being compared directly with English monolinguals. Similarly, I compared both bilingual groups with Arabic monolingual speakers. However, I did not compare the bilinguals directly in their L1 and L2 as this was beyond the scope of the current study. My results therefore speak to L1 attrition and L2 acquisition within the same individuals (in that I examined how they compared with monolingual native speakers of both of their languages). However, my results do not speak to whether there are any differences between L2 English vs. L1-attrited English or L2 Arabic vs.

L1-attrited Arabic because I have not compared bilingual groups directly with each other. My results also do not speak to whether bilinguals show distinct (or merged) patterns in their two languages because I have not compared the English and Arabic of the A-E bilinguals or the English and Arabic of the E-A bilinguals directly. This means that my results speak to the relation between L1 attrition and L2 acquistion within individuals, but not within languages. This will be reserved for future analyses.

The current study tested five hypotheses. I expected differences in the pronunciation of *wh*-words in Arabic and English, with a higher pitch target in Arabic than in English (Hypothesis 1). A direct comparison of *wh*-words between the Arabic and English monolingual control groups confirmed Hypothesis 1 and showed that the pitch pattern for *wh*-words in Arabic had a significantly more inverse U-shaped curvature than that for *wh*-words in English. In other words, monolingual Arabic speakers have a rather steep rise-fall pattern with a high target. By contrast, monolingual English speakers show a rather flat pitch pattern. In addition, I predicted that highly proficient L2 learners would approximate the patterns they find in use by monolinguals (Hypothesis 2). Comparing the pitch patterns of bilinguals' L2 with those of monolinguals, I found that L2 learners of both Arabic and English approximated the respective native patterns when pronouncing *wh*-words, thus supporting Hypothesis 2. Moreover, I expected that, as with other elements of native languages, bilinguals' prosody would show attrition (Hypothesis 3). Interestingly, I found an asymmetric pattern in terms of attrition: While there was evidence for attrition among L1 English bilinguals, there was no such evidence for L1 Arabic bilinguals. Thus, I could only partially confirm Hypothesis 3.

I expected that language use would influence L1 attrition and L2 acquisition (Hypothesis 4). However, I could only partially confirm Hypothesis 4. While I found no evidence that language use influenced L1 attrition, language use did affect L2 acquisition. In particular, learners who self-reported greater use of the L2 showed more native-like patterns in their pronunciation of *wh*-words in that language. Finally, based on the observation that participants in Mennen's (2004) study showed native-like productions in either both their languages or in neither, I tentatively assumed that participants with high phonetic aptitude would show more native-like production of *wh*-words in both their L1 and L2 (Hypothesis 5). However, this hypothesis was not confirmed. Similar to language use, phonetic aptitude did not influence L1 attrition. While phonetic aptitude influenced L2 acquisition, the differences between participants with high vs low phonetic aptitude were such that neither group could clearly be identified as having a more native-like prosody on the *wh*-word than the other. In the following sections, I discuss these results in more detail.

2.7.1 Monolinguals

The data from the monolingual Arabic and English groups confirmed Hypothesis 1 in that the typical pitch pattern for Arabic *wh*-words differed from the typical pitch pattern for English *wh*-words: The *wh*-words produced by the Arabic monolingual group showed a steep rise-fall pattern, which was not found for English. To the best of my knowledge, this is the first empirical investigation of the pitch patterns for *wh*-words in monolingual and bilingual Arabic and English speakers, using carefully controlled materials.

The findings for Arabic confirm what has been suggested in the previous literature. In particular, the Defense Language Institute (1974) proposed that the *wh*-word in Arabic carries a steep rise in pitch, and our findings of a rise-fall pattern are compatible with this description. Furthermore, El Zarka's (1997) corpus evidence suggests that *wh*-words typically carry a pitch accent and yield some of the highest pitch values found in the corpus, again suggesting a rise to a high target on the *wh*-word. She proposed that *wh*-words in Arabic typically carry a ↑H* (L) accent, where the ↑ indicates a raised F0 (El Zarka, 1997). This proposal is compatible with the rise-fall pattern around a central inflection point that I found for Arabic *wh*-words. Note that the present analysis used the existing acoustic signal on the *wh*-word for statistical analysis, and it did not include a detailed ToBI-style manual coding of the data. Owing to the large data set used in the current study, manual ToBI-style coding was not a practical option. As a result, I cannot confirm that the pitch accent on the *wh*-word is indeed a ↑H* (L) accent, but I can say that my data is fully compatible with the proposal that the *wh*-word in Arabic typically carries a ↑H* (L) accent. Future studies – for example, on a subset of my data – would be needed to confirm the particular pitch accent type that *wh*-words in Arabic typically carry.

The findings for English support the claims made by Bartels (2014) but do not confirm Hedberg and Sosa's (2002) corpus results. The *wh*-word in English in Bartels' (2014) theoretical work most often carries an H* accent or is unaccented, which is compatible with the rather flat pitch curve that I found for English in our analysis. By contrast, Hedberg and Sosa (2002) found that *wh*-words in English typically carry an L+H* accent. In this case, I would have expected a steep rise in pitch on the *wh*-word, as is typical for L+H* accents in English (Pierrehumbert & Hirschberg, 1990) and similar to what I found for Arabic. There are several possible reasons why the present data do not confirm Hedberg and Sosa's (2002) corpus data. First, Hedberg and Sosa (2002) extracted *wh*-questions from public affairs discussion programmes, which include rather heated discussions of speakers who widely differ in their political views. As a result, the *wh*-questions analysed in Hedberg and Sosa (2002) may differ widely in pragmatic function from my rather neutral information-seeking *wh*-questions. This may, in turn, influence the prosody on the *wh*-word (see Hedberg et al., 2010, for evidence that pragmatic function strongly influences the nuclear tunes of *wh*-questions in English). For example, Hedberg and Sosa (2010) present as an example the *wh*-question *Why is it going to take a year to put it into place?*, which carries a clear L+H* accent on the *wh*-word. It seems that the speaker who uttered this particular question did not actually seek information but was rather asserting that it should *not* take a year to put 'it' into place. Thus, the speaker seems to be incredulous that 'it' would take a year to put into place, and the L+H* on the *wh*-word here may contribute to conveying this incredulity. Note that this particular question could quite felicitously be asked with an H* or no pitch accent on the *wh*-word and such a pattern would be quite compatible with the kind of information-seeking questions that I used in the current study. It may also be the case that the *wh*-question word *why* is more likely to carry a L+H* accent in English than other *wh*-question words. My study did not include any instances of *why*, and I don't know how many *wh*-questions with *why* were included in the 35 *wh*-questions that Hedberg and Sosa (2002) analysed. Overall, pragmatic differences or the choice of *wh*-word may be driving the conflicting results between Hedberg and Sosa's (2002) and my current findings.

As mentioned above, I did not conduct a detailed, manual ToBI-style analysis of the pitch pattern for English and Arabic *wh*-words because it would have been prohibitively time-consuming for the large amount of data that I analysed. Instead, I compared the acoustic signal in terms of F0 contour for *wh*-words in English and Arabic directly, using a statistical analysis that allows us to compare the slope and curvature of curves. This method still involves some manual coding because the beginning and end of the *wh*-words need to be marked. However, the manual workload is greatly reduced compared to manual ToBI-style analysis. I believe that this approach is particularly useful for large data sets when the research question does not revolve around which particular pitch accents occur in a certain position, but rather whether differences exist among particular items in terms of how prosody is realised in a certain position. In particular, the GCA that I used here shows that the average pitch curve on the *wh*-word is significantly steeper and has significantly more inverse U-shaped curvature in Arabic than in English. However, the results from the GCA fail to show from which pitch accent distribution these differences derive. For example, the differences in pitch pattern that I find for Arabic and English *wh*-words could be due to Arabic and English using different pitch accents to mark the *wh*-word. For example, my results are compatible with Arabic typically using a ↑H* (L) accent, as proposed by El Zarka (1997), and English typically using an H* accent or no accent, as proposed by Bartels (2014). Alternatively, the differences could be due to Arabic and English using the same or similar pitch accents to mark the *wh*-words, but differing in how frequently each language uses these same pitch accents. For example, my results are compatible with Arabic typically using a ↑H* (L) accent, and sometimes though rarely an H* accent, and English typically using an H* accent and rarely an L+H* accent. Thus, while the current analysis can detect differences in the signal, it may not always be clear from where these differences derive.

In terms of the analysis, I would like to highlight that previous analyses that have examined the prosody in L1 attrition have mostly considered pitch peak alignment (e.g. Mennen, 1999; de Leeuw, 2008). Such alignment is a valuable measure when two languages or varieties display the same overall pitch accent pattern, but may differ in the alignment, typically, of a peak. However, in the current study, such an analysis was not possible for two reasons: Owing to the little amount of previous work on the prosody of the *wh*-word in *wh*-questions in both English and Arabic, I did not know with enough confidence which pitch patterns to expect for the *wh*-words. Thus, looking for differences in pitch peak timing would have been premature. In addition, the previous literature suggested that there was a possibility that the *wh*-word in English may frequently be unaccented. In this case, the location of a pitch peak could not be determined. I therefore suggest that the current analysis may be especially useful for exploring L2 prosodic acquisition and L1 prosodic attrition when little information is known about the prosodic patterns that may be found in the compared varieties and when the compared varieties may differ, not in timing, but in the overall shape of the pitch contour.

2.7.2 Bilinguals: L2 Acquisition

I found that both groups of highly advanced late L2 bilinguals (A-E and E-A bilinguals) approximated native speech when producing *wh*-words, thus supporting Hypothesis 2. In particular, I found no significant differences in term of inverse U-shaped curvature between the monolinguals and the bilinguals L2 speech, which suggests that the bilinguals in this study managed to approximate the prosody of *wh*-words in their L2. More specifically, even though the bilinguals' L2 patterns differed from those of monolinguals, I found no evidence that the differences I did find were influenced by the bilinguals' L1. It is probably unsurprising that highly proficient bilinguals approximate native norms, especially those who have been in the host country where the L2 is spoken for a long time.

The current results of the bilinguals confirmed LILt, which holds that, as learners increase their experience of the L2, their intonation parameters will more closely approximate the norms of the L2. In more detail, both A-E and E-A bilinguals, who are highly advanced late L2 bilinguals with prolonged experience of the L2, approximated native speech when producing *wh*-words.

However, looking at the findings of the current research through the lens of cross-linguistic interference, they are somewhat surprising. In particular, I found no evidence that bilinguals' L1 interfered with their L2 productions of *wh*-words. This finding differs from the results of previous studies by Mennen (1999), Ramirez and Romero (2005) and Willems (1982), which suggested that the L1 influenced the prosody of the L2.

The present results differed from Mennen's (1999) findings, even though the participants in both studies were all near-native highly proficient bilinguals. There are several possible reasons for this discrepancy. In the current study, I looked at the global pitch pattern for *wh*-words in questions rather than pitch-peak alignment, which was studied by Mennen. It is possible that pitch-peak alignment is a more fine-grained measure and can detect smaller differences than my more global measure. In addition, Mennen (1999) analysed the speakers in her research individually, whereas I did an analysis of the signals of the curves for the participants as a group, which means that the sample in this study might contain individual participants who show L1 interference, even though I found no evidence for this at the group level.

Furthermore, my study differed from Ramirez and Romero's (2005) and Willems' (1982) study, which showed an effect of the L1 on speakers' L2 productions. This effect may be because the participants in the current study were considered to be near-native highly proficient bilinguals, while Ramirez and Romero (2005) and Willems (1982) tested learners of English who were not advanced bilinguals.

My results partially confirmed Hypothesis 4. More specifically, E-A bilinguals with high L2 use produced more native-like patterns than did E-A bilinguals with low L2 use. Moreover, A-E bilinguals with high L2 use also seemed to approximate the native pattern more closely than did A-E bilinguals with low L2 use. Thus, overall, participants with more L2 language use showed more native-like production of their L2 than did participants with less L2 language use. These results are in line with previous research by Purcell and Suter (1980), who identified that living with English native speakers (greater L2 use) affected the degree of foreign accent in an L2. Hence, more frequent use of the L2 with native speakers relates to more native-like speech. Similarly, Thompson's (1991) study showed that a significant correlation exists between L2 English use and the degree of L2 foreign accent. The current study is also in line with Hopp and Schmid's (2010) work, which showed that the more L2 learners of German spoke German with a partner or with native speakers, the better their pronunciation. These previous studies suggest that more L2 use and in particular L2 use in the home and with native speakers is conducive to native-like pronunciation in the L2, which is also in line with the current results. Participants with high L2 use in the current study used their L2 with friends, partners, family and work. The work environment, especially for the E-A bilinguals, typically involved mostly interactions with native speakers of the bilinguals' L2. By contrast, participants with low L2 use tended to use the L2 at work and with friends. They were thus less likely to use the L2 in the home with family members or partners. In addition, some of the friends with whom they used the L2 were also non-native speakers.

I did not confirm Hypothesis 5. While phonetic aptitude did influence participants' productions, there was no evidence that the differences in productions between participants with high and low aptitude related to the nativeness of the productions. In other words, the present study found that neither group of bilinguals appeared more or less native-like in their L2. This finding contrasts with previous studies, such as Jilka et al. (2008) and Hopp and Schmid (2010). Jilka et al.'s (2008) work found that phonetic aptitude influenced the prosodic pattern of German-English bilinguals, which may be because Jilka et al. (2008) conducted a large-scale project to examine pronunciation talent in detail, whereas the present study was done on a somewhat smaller scale. In addition, the tests I used to measure aptitude differed from those used by Jilka et al. (2008). I used three tests that were adapted from the MLAT and PLAB, and that either tested participants' knowledge of sounds, such as phonetic representations, or pattern recognition skills. Both Jilka et al. (2008) and the current study used “Spelling Cues” and “Paired Associates” tests, which were similar, but not identical. In addition, Jilka et al. (2008) used other tests from the MLAT, including, for example, a test that measures participants’ comprehension of syntactic structures. In contrast, I used the Sound Discrimination test to distinguish pitch, orality and nasality in a language with which participants were unfamiliar (PLAB; Pimsleur, 1968), which is not used by Jilka et al. (2008). Overall, therefore, the tests used in Jilka et al. (2008) covered a wider set of skills than the tests used here, which are not entirely, but more closely focused on sounds. It is thus possible that my results are not in line with Jilka et al.’s (2008) because I used different tests to measure aptitude. Specifically, while the tests in Jilka et al. (2008) measured what could most closely be described as language aptitude, I measured skills that loosely involve skills related to sounds and therefore more closely, but not entirely, relate to phonetic aptitude such as ‘sound discrimination test’.

Notably, my results are compatible with Suter’s (1976) findings that musical ability (i.e. aptitude for sounds and melodies) did not affect the degree of L2 foreign accent, and with Flege et al.’s (1999) findings that their measure of sound processing ability contributed little to an L2 foreign accent. Overall, these results suggest that some skills and talents may positively relate to better L2 pronunciation, but it is not yet clear which skills and talents these are or how they can be measured. Further studies are therefore needed to determine which particular aptitude tests relate to more native-like pronunciation and what kinds of talents these aptitude tests actually measure.

2.7.3 Bilinguals: L1 Attrition

The data in this study partially confirmed Hypothesis 3 in that I found evidence for L1 attrition only in the L1 English bilinguals, but not in the L1 Arabic bilinguals. Thus, the results from the E-A bilinguals, but not the A-E bilinguals, are in line with previous studies showing attrition in the L1 (Mennen, 2007; Mennen et al., 2007; de Leeuw, 2008; de Leeuw et al., 2012; Hopp and Schmid, 2010). The results partially confirmed LILt’s assumption that there is a bidirectional influence between L1 and L2 in the L1 English bilinguals. As LILt did not make any specific prediction as to when you would or would not find attrition in intonation, it is not entirely clear how the assymetric attrition pattern found here relates to LILt. Specifically, since LILt assumes that bidirectional influences between the L1 and L2 can occur, the results from both the E-A bilinguals and the A-E bilinguals are generally compatible with LILt. However, LILt would not currently serve as an explanatory framework for why L1 attrition was only found in E-A bilinguals, but not in A-E bilinguals.

There are several possible reasons for the discrepant results of the E-A and A-E bilinguals. In particular, my results are compatible with Eckman's Markedness Differential Hypothesis (MDH; Eckman 1987, p, 1), which states that "when two languages differ, marked structures are more difficult to acquire than unmarked structures". According to Zerbian (2015), English sentence prosody is more marked than Arabic sentence prosody. "Prosodic focus marking (e.g. through pitch accent placement) and prosodic givenness marking (e.g. through deaccentuation) are thus two independent factors that can each contribute to pragmatic constraints in sentence prosody" (Zerbian, 2015, p.15). Zerbian (2015) states that all languages that have givenness marking also have focus marking, but some languages only have focus marking and no givenness marking. In terms of the MDH, this means that languages that have givenness marking in their sentence prosody (English) are more marked than are languages that do not have givenness making (Arabic), which means that English sentence prosody is predicted to be more difficult to learn than Arabic sentence prosody. It is therefore possible that the English prosody of *wh*-words is harder to acquire than the Arabic prosody of *wh*-words. If something is harder to acquire in the L2, then it typically takes longer to acquire and may never be fully acquired. It thus makes sense that something that is harder to acquire would also be less likely to influence the L1. In particular, an aspect of the L2 that is harder to acquire would require more L2 exposure than an aspect of the L2 that is easier to acquire to cause attrition in the L1. The A-E bilinguals in our study may show no evidence of L1 attrition in their production of *wh*-words because English sentence prosody is sufficiently difficult to acquire and thus unlikely to influence the L1. Similarly, if something is easier to acquire in the L2, then it is typically acquired rather quickly and may be more likely to influence the L1. In particular, an aspect of the L2 that is easier to acquire would require less L2 exposure than an aspect of the L2 that is harder to acquire to cause attrition in the L1. The E-A bilinguals in the present study may thus show evidence of L1 attrition in their production of *wh*-words because Arabic sentence prosody is comparatively easier to acquire and thus more likely to influence the L1.

However, it is possible to explain the current results without reference to a particular theory, such as the MDH. Specifically, I could alternatively explain the current findings in terms of acoustic salience or acoustic prominence. The Arabic *wh*-word is acoustically more salient and prominent because it involves a large pitch excursion and thus a noticeable change in pitch during the production of the *wh*-word, which makes it perceptually salient. By contrast, the English pattern of *wh*-words is acoustically less salient or prominent because it does not involve much of a pitch excursion. As a result, listeners may more easily pick up the prosodic pattern of Arabic *wh*-words compared to English *wh*-words. A pattern that is more easily perceived and picked up in the L2 may also be more likely to show an influence on one's L1 and thus contribute to L1 attrition. By the same logic, an aspect of the speech stream that is acoustically very salient in the L1 may be resistant to attrition from an L2 where this aspect of the speech stream is less salient. Thus, an explanation based on acoustic salience or acoustic prominence can also explain why I found L1 attrition in the E-A bilinguals, but not in the A-E bilinguals.

The attrition data did not support Hypotheses 4 and 5. In particular, the L1 Arabic bilinguals showed no L1 attrition to begin with, and neither L2 language use nor phonetic aptitude modulated the L1 English bilinguals' productions in a way that one group could be considered to be more native like than the other. Thus, I found no evidence that language use or phonetic aptitude influences L1 attrition. It is possible that language use and phonetic aptitude did not affect L1 attrition because the L1, despite still being malleable in terms of phonetics in adulthood, is relatively more stable than the L2. In other words, it is likely that pronunciation differences across individuals are smaller for L1-attrited speech than L2 speech. This would be in line with the current results as I found an asymmetry between L2 acquisition and L1 attrition in that differences in terms of language use (but not aptitude) modulated L2 acquisition, but not L1 attrittion, suggesting that the latter is less likely to be modulated by factors such as language use or aptitude.

2.7.4 Conclusion

To conclude, the present study confirmed that the prosodic patterns of *wh*-words in Arabic and English differ reliably. In line with LILt, I have found that the E-A and A-E bilinguals approximate the norms of the L2. Language use seemed to contribute to the acquisition of the prosody of *wh*-words in the L2. In line with Markedness Theory or the notion of acoustic salience, I found L1 attrition in L1 English bilinguals but not in L1 Arabic bilinguals. In addition, while the results of both bilingual groups are generally compatible with LILt, the theory does not provide an explanation for why only E-A, but not A-E, bilinguals showed L1 attrition. Overall, the current results suggest that individual differences and aspects of the particular languages interact in the acquisition and attrition of prosody.

Chapter Three (Study II)

L2 acquisition and L1 attrition of vowels in highly-proficient late bilinguals: Exploring the role of phonetic aptitude and language use

Abstract

Comparatively little research has examined second-language (L2) acquisition and first-language (L1) attrition of shared vowels, and few studies have examined the effect of individual differences in late bilinguals in terms of L2 acquisition and L1 attrition of segments. The present chapter investigates L2 acquisition and L1 attrition in the segmental productions of late E-A and A-E bilinguals. In addition, I investigate whether late L2 learners' phonetic aptitude and language use affect how target-like learners are in the shared vowels of both Arabic and English. I present data from 60 participants divided into four groups: 15 monolingual Arabic speakers, 15 monolingual English speakers, 15 E-A bilinguals and 15 A-E bilinguals. The bilinguals had been living in their L2 environment for an average of 20 years. All participants read a list of words and sentences containing different vowels, and bilinguals read them in both Arabic and English. In addition, bilinguals completed an L2 proficiency test, a phonetic aptitude (talent) test and a language-background questionnaire. Vowel duration was measured from the approximate beginning of the selected vowel to the end of the vowel, and a Praat script extracted F1, F2 and F3 for all the selected vowels spoken in isolation and in a carrier phrase. Vowels were normalized with the Bark Difference Method (Syrdal & Gopal, 1986). The results from the monolingual groups revealed differences in some of the shared vowels in Arabic and English. Moreover, I found L1 attrition among both the A-E bilinguals and the E-A bilinguals. Finally, I found no evidence that language use or phonetic aptitude influenced the segments of A-E and E-A bilinguals in L1 attrition and L2 acquisition.

Keywords

Bilinguals, second-language (L2) acquisition, first-language (L1) attrition, shared vowels, first formant (F1), second formant (F2), language use, phonetic aptitude (talent)

3.1 Introduction

Many adult late bilinguals are clearly identifiable as non-native speakers when speaking in their second language (L2) due to the influence of certain properties of their L1 on their L2 (Scovel, 1969; Brennan et al., 1975; Flege, 1980, 1981). Furthermore, some studies have investigated why the L1 affects the L2 (Flege, 1999; Moyer, 2009). However, in recent years, a phenomenon called L1 attrition has attracted the interest of linguists and psychologists. Specifically, it has been found that the L2 influences the L1 and that late L2 learners are sometimes found to differ in the production of their L1 compared to native speakers (Mennen, 2004). In this chapter, as in previous chapters, L1 attrition is defined as the "non-pathological, non-age related structural loss of a first language within a consecutive bilingual" (de Leeuw, 2008, p.10). Therefore, L1 attrition applies to bilinguals who, at some point in time, begin to lose full command of their L1. This chapter investigates L2 acquisition and L1 attrition in the segmental productions of late bilinguals of Arabic and English to evaluate the possible influence of the L1 on the L2 and of the L2 on the L1 in their phonetic systems.

An important theory in this regard, and one which is highly relevant to the present chapter, is the SLM proposed and later refined by Flege and colleagues (Flege & Hillenbrand, 1984; Flege, 1987a, 1995; Flege et al., 2003). As I mentioned before in Chapter One, section1.2.3, the aim of the SLM is to explain how learners acquire speech sounds in the L2. It also contributes to our understanding of how this process in turn affects the L1. Many studies of bilingual migrants have explored what happens when speakers encounter sounds in the L2 that are different from those in the L1. Some previous studies have focused on the period of initial contact (Flege et al., 2003), while others have focused on the ultimate attainment of learners (Bongaerts et al., 1997), and still others have attempted to trace the development that occurs in phonological perception and production over long periods (de Bot & Clyne, 1997; Cook, 2003). As mentioned previously (in section 1.2.3), the SLM model highlights the importance of the degree of difference between L1 and L2 sounds and predicts that L2 sounds that are similar to L1 sounds are assimilated into an existing L1 category, whereas L2 sounds that are dissimilar to L1 sounds are incorporated as a new category. The PAM-L2 model makes similar predictions. Specifically, PAM-L2 proposes that L2 segments that are similar to the L1 will be 'perceptually assimilated' to the categories of the L1 and deviances from L2 production are likely to occur. In contrast, the L2LP model makes the opposite prediction in suggesting that dissimilar segments are the ones that pose particular difficulties.

What is clear from the empirical studies on this subject is that “all else being equal, early bilinguals will be more likely to establish new phonetic categories for L2 speech than late bilinguals will be” (Flege et al., 2003, p.469). However, it is less clear what happens in situations with prolonged and frequent contact L2 contact, possibly in combination with ongoing frequent L1 contact. It could be, for example, that frequent contact with L1 speakers in an L2 environment promotes the formation and retention of hybrid categories rather than of distinct L1 or L2 categories. The present chapter investigates late consecutive A-E and E-A bilinguals’ L2 acquisition and L1 attrition in the production of vowels, specifically long and short vowels that are similar in Arabic and English. I investigate similar vowels because the SLM and PAM-L2 theories predict that bilinguals would have merged rather than separate categories for similar vowels. I would thus expect that late bilinguals produce these similar L2 vowels as they do in their L1. In addition, if merged categories shift in the direction of the L2 after prolonged L2 exposure, I would expect L1 attrition for vowels that are similar in English and Arabic. I investigate the production of vowels by monolingual native speakers of both languages and the relationship between L1 attrition and L2 acquisition in highly fluent, late consecutive A-E and E-A bilinguals. In addition, I explore how the amount of L2 language use and phonetic aptitude affect both L1 attrition and L2 acquisition in their phonetic systems. The present chapter adds to the existing literature on vowel productions in L2 acquisition and L1 attrition.

3.2 Literature Review

3.2.1 Acquisition of L2 Vowels

Previous studies have shown that the L2 vowel productions of late bilinguals frequently differ from those of native monolingual speakers (Major, 1987; Busà, 1992, 1995; Flege, 1992a, 1992b; Munro, 1993; Jun & Cowie, 1994; Flege et al., 1997a). Several researchers have examined bilinguals’ ability to perceive L2 vowels (Flege, 1992a, 1992b; Best et al., 1996; Flege et al., 1997a, 1999). In addition, a large number of studies have investigated late bilinguals’ ability to produce L2 vowels (Flege, 1987, 1992; Munro, 1993; Moyer, 2009). The findings of these studies show that early bilinguals who learned the L2 during childhood were able to perceive and produce the vowels of the L2 more accurately than could late bilinguals who learned English after the onset of puberty. For example, Munro et al. (1996) investigated the effect of age on the production of English vowels by native Italian second-language learners of English, and surveyed 240 speakers. The participants varied in terms of their AOA in the host country (Canada) and LOR. The authors obtained production samples via a delayed repetition task in which the participants repeated English words. Native English listeners assessed the participants' accuracy when producing vowels by rating the vowel sounds on a five-point Likert scale. Age was indeed a factor; however, the researchers also noted considerable variation in the ratings that the listeners gave to the utterances in the L2, and variation was also evident in the performances of different vowels. Most of the Italian participants who arrived in the host country before the age of 12 produced /ɚ/ as would a native speaker; however, a small number of the participants who had arrived in Canada later also produced this sound accurately.

There are several studies (Bongerts et al., 1997; Nikolov 2000; Bellingham; 2000; Neufeld, 2001) that contradict the Critical Period Hypothesis (CPH), in that they state that adults do have the potential to achieve a native-sounding accent. Nokolov (2000) found 11 out of 33 participants were sometimes mistaken for native speakers when using their L2. This offers a counter-argument to that presented by the CPH. Her survey displayed that "these successful language learners want to sound like natives, they share intrinsic motivation in the target language which is often part of their profession, or they are integratively motivated[...]. They work on the development of their language proficiency consciously and actively through finding chances for communicating with speakers of the target language, reading and listening extensively" (Nokolov, 2000, p. 122). Some researchers have therefore questioned whether it is possible to apply the CPH more broadly, and to the acquisition of L2 phonology in particular (Flege, 1987; Abu-Rabia & Kehat, 2004). For example, Moyer (2009) argued that although age is a factor in L2 acquisition, it is not always the most important factor. As mentioned previously, highly motivated adults who are immersed in the L2 environment and supported by good instructors can achieve very high levels of speech performance.

Some studies on the acquisition of L2 vowels have found support for the SLM and PAM-L2 hypotheses, which suggest that sounds that are similar, but not identical, in the L2 and L1 are the most difficult to learn because they are perceived to be similar. For example, Flege and Hillenbrand (1984) investigated why many people who learn a foreign language tend to speak with a recognisable foreign accent. The authors hypothesised that certain factors limited the human ability to acquire a second sound system after L1 acquisition is complete. It appears that when L2 phones are similar to L1 phones, L1 phones tend to be used in place of L2 phones. When new L2 phones are encountered, these may be acquired accurately, or sometimes incorrectly identified as similar but not identical L1 phones. Flege and Hillenbrand (1984) designed three speaking tasks to examine how American L1 English speakers produced the L2 French sounds /tu/ (*tous*) and /ty/ (*tu*) compared to French L1 speakers. As would be predicted by the SLM, L1 American speakers produced the /y/, which is different from any English vowel, indistinguishably from French L1 speakers, but they did not produce the /u/ in the same way as the native French speakers. Rather, they merged the French /u/ with their similar American English /u/.

Additionally, a study by Flege (1987) revealed that adult L2 learners with sufficient L2 exposure were able to produce some vowels in the L2 accurately. In line with Flege and Hillenbrand's (1984) results, native English speakers who were highly experienced L2 speakers of French produced /y/ in French accurately, while their /u/ sound in French remained as they produced it in English. In particular, English-French bilinguals produced significantly higher F2 values of /u/ in French than did monolingual French speakers, which suggests that the L1 English speakers considered the French /u/ to correspond to the /u/ in English. By contrast, English has no vowel similar to the French /y/, leading native English bilinguals to create a new category and produce the novel vowel accurately. A further observation in line with the SLM hypothesis is that experienced or highly fluent late L2 learners can produce certain L2 vowels accurately, particularly "when the L2 vowels are located in a portion of vowel space that is not occupied by an L1 vowel" (cf. Flege, 1987, 1992b; Bohn & Flege, 1992; Ingram & Park, 1997; Flege & Mackay 2004, p.2).

Flege's (1992) study supported the SLM hypothesis. The study tested the predictions of the SLM by examining the L2 intelligibility of the similar vowels /iː/, /ʌ/ and /ʊ/ in productions by native Dutch L2 English learners. The study investigated 50 Dutch university students who had started learning English as an L2 after the age of 12 and compared their productions to native English speakers from different dialect regions to assess the intelligibility of the speakers. Flege (1992) found no significant difference in intelligibility: the /iː/, /ʌ/ and /ʊ/ vowels were equally intelligible when produced by native Dutch speakers with mild, moderate and strong foreign accents. Nevertheless, the Dutch speakers did not vary from English native speakers when producing the vowels /iː/, /ʌ/ and /ʊ/, and an improvement in the production of the similar vowel /ɒ/ was noted with increased L2 experience. The mildly and moderately accented Dutch speakers showed more accurate production of /ɒ/ than did the strongly accented Dutch speakers. The results revealed "that the phonetic relationship between vowels in the L2 and in the native language (L1) is an important determinant of the intelligibility of L2 vowels" (Flege, 1992, p.216). In addition, the study showed inherently varying intelligibility across the different vowels. Moreover, it was demonstrated that the native listeners' dialects played a role in determining the accuracy of spoken vowels in an L2. For example, the native British English productions of /ɪ/ and /i/ were recognised at near-perfect rates, but their /æ/ and /ʌ/ were relatively poorly identified. "The greatest difference between the native and Dutch talkers was for /ʊ/ and /a/. The Dutch talkers' vowels were less intelligible than the native talkers' because their /ʊ/ was heard as /ʊ/ and their /a/ was heard as /ʌ/" (Flege, 1992, p.205).

By contrast, Wu and Shih's (2012) study did not support the predictions of the SLM. Wu and Shih (2012) examined learners' foreign accent by comparing different, identical, similar but not identical, and different vowels in the productions of Chinese native speakers, Chinese heritage speakers and Chinese L2 learners. Undergraduate students in Taiwan who were Mandarin native listeners (n = 43) rated the accentedness of learners' productions "to measure the perceptual distance of speech between speakers and listeners, such as dialect accent and foreign accent" (p.2). Data from a large corpus of spontaneous speech of Chinese learners showed that pronunciation, which had been affected by an individual's L1, is likely to be perceived as having a 'foreign' sounding accent. Therefore, a strong relationship is evident between accent and pronunciation in the rating results. However, the results of this study failed to support the predictions of the SLM since "the L2 learners did succeed in learning similar vowels and had problems in learning new vowels, as well as identical vowels" (Wu & Shih, 2012, p.4).

3.2.2 Attrition of L1 Vowels

The majority of the literature on L2 acquisition has focused on the formation of the L2 as influenced by the L1. Because the L1 is considered to be a more stable linguistic system than the L2, it is deemed more likely that the L1 will influence the L2 than it is that the L2 will influence the L1. As Schmid and Köpke (2007, p.4) stated, "The phonological system of a mature L1 is probably so stable that it is impervious to L2 influence". Recently, more research has aimed to analyse the influence of the L2 on the L1 (Chang, 2012; de Leeuw et al., 2013).

Literature on L1 phonological attrition is scarce (Chang, 2012, 2013; Mayr et al., 2012; de Leeuw et al., 2013). However, there is general evidence for attrition in the L1 in the phonetic domain at the segmental level (Flege & Hillenbrand, 1984; Flege, 1987; Major, 1992; Sancier & Fowler, 1997; Ventureyra et al., 2004; Chang, 2012; Mayr et al., 2012; de Leeuw et al., 2013; Ulbrich & Ordin, 2014; Kartushina et al., 2016 ; Bergmann et al., 2016 ; Cabrelli et al., 2019) and at the suprasegmental level (Mennen, 2004; de Leeuw et al., 2012). Various studies have suggested a bidirectional influence of L1 and L2 segmental sound systems that is measurable at a fine-grained level of detail when languages differ in their vowels (Guion, 2003; Chang, 2012, 2013; Mayr et al., 2012). Examples of L1 attrition of this kind have been analysed using many language pairs. An important study by Mayr et al. (2012) compared the level of L1 attrition in 62-year-old monozygotic twin sisters who were bilingual in Dutch and English, where one of the twins had moved to the UK at the age of 32. Both twins used Dutch and English daily, but the twin (MZ) who moved to the UK used English more often than she used Dutch, while her counterpart (TZ) in the Netherlands did the opposite. In general, all the vowels of Southern Standard British English (SSBE) are more open than the Standard Dutch equivalents (Deterding, 1997; Adank et al., 2004; Hawkins & Midgley, 2005). Studies of twins have the advantage of allowing researchers to exclude certain variables, thus isolating the effects of a smaller range of variables. Flege's (1995) SLM suggested that cross-linguistically similar L1 and L2 categories will result in merged L1–L2 representations. Consistent with the SLM, MZ did not produce a cross-linguistic difference between some of her similar English and Dutch vowels. One example is the vowel /ɛ/. Mayr et al. (2012, p.696) suggested that the vowel /ɛ/ in MZ's Dutch "may have been 'pulled' towards a more open, and thus more English-like position due to interlingual identification with English /ɛ/".

> "The same mechanism could then be responsible for changes to other L1 categories. Thus, according to this account, the observed shift in F1 across the various Dutch categories could be the result of a series of unconnected changes affecting pairs of L1 and L2 vowels" (Mayr et al., 2012, p.698).

Overall, the authors observed systematic differences between the two speakers that affected some phonemes and not others, suggesting that attrition may not affect all areas of pronunciation equally (Mayr et al., 2012). The results showed fundamentally different vowel spaces for both sisters; in addition, the result for MZ's vowels in Dutch and English revealed no contrast for some of MZ's English and Dutch vowels. Overall, Mayr et al. (2012) suggested that attrition may operate on a system-wide basis rather than item-by-item basis. It needs to be mentioned though that the current study cannot distinguish between L1 system-wide vs. item-by-item attrition because the present study is looking at a subset of vowels of each of the languages involved.

In a similar vein, Bergmann et al. (2016) studied a group of adult German immigrants in English-speaking North America, focussing on L1 attrition. The participants were observed in comparison to monolingual Germans. Participants were rated for how native they sounded. As expected, those who had emigrated sounded less native in German than the control group. Indeed, the longer the participants had spent away from their home country (and the more they used English in everyday life), the less native they sounded. An additional formant analysis for the four German phonemes /aː/, /ɛ/, /ɔ/ and /l/ was conducted for two groups of bilinguals: the more native-sounding ones, and the less native-sounding ones, based on the previous foreign accent ratings. Although it was hypothesised that L1 attrition in the direction of the L2 would be stronger in the less native-sounding group, both groups' pronunciation was equally L2-like. Furthermore, only a weak link between the formant values and the ratings was established, suggesting that there were other segments, and possibly suprasegmentals, which made the participants sound non-native.

L1 attrition in the phonetic domain can also occur due to polarisation, which "serves to enhance the contrast between categories in a phonetic system" (Flege, 1991a, p.280), for example, when the sound (i.e. /ʊ/) moves away from both the typical L1 and L2 pronunciations. According to Mayr et al. (2012, p.688), "Polarisation has also been observed in the context of bilingual vowel systems". For example, Guion (2003) examined the production of the Spanish vowels /i/, /e/, /a/, /o/ and /ʊ/ and the Quichua vowels /ɪ/, /a/ and /ʊ/ by L1 Quichua L2 Spanish bilinguals. Four types of bilinguals were examined, namely simultaneous, early, mid and late bilinguals. The groups that managed to keep the categories of the L1 and the L2 distinct were the simultaneous bilinguals, early bilinguals and some mid bilinguals. The group of late bilinguals produced vowels in both languages (L1 Quichua and L2 Spanish) with L1 categories. Notably, the participants who acquired the L2 vowels produced their L1 Quichua vowels higher in the vowel space than did the participants who had not acquired the L2 vowels. This shift could be argued to be the result of the participants' attempts to distinguish perceptually between the two languages (cf. Guion, 1986; Lindblom, 1998).

L1 attrition does not necessarily involve polarisation, but it can also involve a shift of L1 vowels in the direction of the L2. For example, Chang (2012, 2013) described transfer in the vowel system of the L1 of late English–Korean bilinguals in the direction of the L2 system. The longitudinal research by Chang (2012) is based on the hypothesis that changes in L1 phonetic categories are not limited to highly advanced L2 learners, but they can also occur in the early stages of L2 learning. This is evident in Chang's (2012) findings, which examined the speech of learners of Korean with L1 American English in two production experiments. The control groups consisted predominantly of monolingual speakers of Korean. The participants read English and Korean monosyllabic words for which the plosive and vowel productions were measured. The longitudinal results showed a significant phonetic drift of the target sounds in the participants' L1 to L2 phonetics after six weeks. Chang's (2012) finding lends credence to the idea that this attrition in the L1 is due to recent L2 experience as opposed to the belief that L1 attrition is the result of using the L1 less frequently, as most research has suggested.

Another experiment in phonetic production supports the hypothesis that learners of an L2 are more easily influenced by more recent exposure to the L2 with regard to pronunciation in the L1. Chang (2013) examined productions of L1 American English speakers who were learning Korean. Participants were divided into an experienced group and an inexperienced group. The study investigated changes in the VOT (which is not discussed in this chapter) and vowel production, and placed the participants in the same learning environment – an intensive course in Korea lasting six weeks. The study tested the phonetic drift from the L1 to the L2 by comparing the two groups. Chang's (2013) hypothesis was that the more experienced group would have less of a drift than would the inexperienced group. The results, measured via acoustical speech analysis, revealed that L1 vowels lengthened considerably in both groups of speakers, but that this effect was not as prominent amongst the experienced learners. Inexperienced learners showed more of an L1 vowel drift towards the longer L2 vowels than did the group of more experienced learners.

Similarly, Chang's (2012, 2013) research described a change in the vowel system from the L1 of late English–Korean bilinguals in the direction of the L2 system. Specifically, the reason for the shift was not because of the need for a greater distinction between the categories in the L1 and in the L2, as in Guion's (2003) study, but rather because of assimilatory processes, which is similar to the findings of Flege (1987) and Major (1992). Chang (2012) debated whether the difference between his and Guion's (2003) studies may have been a result of the differences in the onset age of learning the L2 because Chang's (2012) participants were late bilinguals and Guion's (2003) participants were early bilinguals. Chang's study also has implications for the current study as the current participants are very experienced bilinguals. The L2 is therefore not novel to them and the increased attrition affect that Guion found for inexperienced learners should not occur in the current participants. Instead, the L2 should influence the L1 relatively less.

In summary, several studies have suggested a bidirectional influence of L1 and L2 segmental sound systems that is measurable at a fine-grained level of detail when languages differ in their vowels and is evident in more global measures such as bilinguals' perceived foreign accents.

3.2.3 Factors that Influence L1 Attrition and L2 Acquisition

Chapter Two explored factors that influence L2 acquisition and the degree of L1 attrition in the phonetic domain – in particular, language use and phonetic aptitude. These factors influenced L2 acquisition and the degree of L1 attrition in the prosody of *wh*-words in A-E and E-A bilinguals (see Chapter Two for more details). Hence, these factors may have an effect or play a role in A-E and E-A bilinguals' productions of shared vowels.

The Role of Language Use in L2 Acquisition and L1 Attrition

There are a number of studies that have shown that language use affects the strength of a foreign accent overall, (see Chapter Two, section 2.2.3.1, for more information). This section presents some studies that have specifically looked at language use and the production of vowels. Previous research has shown that increases in experience with the L2 (greater amount of L2 use) might support or actively encourage L2 learning. To tell native from non-native sounds, frequent exposure to one's L2 is vital (Flege, 1991; Ingram & Park, 1997), as the speaker will be better able to perceive and therefore produce the L2 sounds in a natural setting (e.g. Flege & Bohn, 1997; Flege et al.,1997).

The accuracy of vowel production from those who started learning an L2 in adulthood is not consistent within that group. Major (1987) studied vowel production of students in Brazil with a 'relatively good' to 'poor' pronunciation of English. Generally, the stronger the foreign accent, the less intelligible are the /æ/ phonemes that students produce. Flege (1992a) found the same pattern of results for L1 Dutch speakers of English. Thus, good pronunciation overall is reflected in the production of individual vowels. However, this begs the question of whether improvements in vowel production in adults depend on long-term natural exposure to an L2 or whether some inherent propensity for L2 acquisition is required.

Jun and Cowie (1994) examined Korean adults who had been residing in the US for either 1-5 or 26-31 years. The results showed that the Koreans who had spent longer in the US produced the English /ɪ/ sound more native-like than the other group. One might conclude from this that longer exposure to an L2 comes hand-in-hand with better production of the L2 sounds. Similarly, Munro (1993) found that Arabic-speaking adults living in the US for 5.7 years produced English vowels, even those without a direct phonetic counterpart in Arabic, in a different way to native English speakers.

Munro's (1993) study adds to this in that no significant correlation was found between time spent in a certain country and any improvement in L2 vowel production. To interpret these findings, one must first know the rate at which adults acquire L2 vowels. Two of the participants had lived in the US for less than two-and-a-half years. The Arabic-speaking participants may still have been in the process of improving their pronunciation at the time of the study. Although, if it is true that the vast majority of L2 acquisition takes place early-on in someone's exposure to the L2 (within two years, perhaps), it may not have been possible to observe a relationship between the LOR in a country and vowel production accuracy because the participants would have reached their full potential for L2 vowel learning.

Last, a study by Bohn and Flege (1992) examined whether adult L2 learners could produce and perceive a new vowel category in an L2. They aimed to test the effect of L2 experience on category formation for new vowels and to study the relation between perception and production in the acquisition of a new vowel category (Flege et al., 1996). While Bohn and Flege (1992) addressed both production and perception, the current study focuses solely on production. Bohn and Flege (1992) examined native German learners of English in the production of English /æ/ compared to native English monolinguals. The German participants were divided into two groups: experienced participants who had lived in the US for five years (mean length of stay: 7.5 years) and inexperienced participants who had arrived in the US less than a year earlier (mean length of stay 0.6 years). The mean age of both groups was 28 and 33 years, and the participants had studied English at school for the same number of years. Acoustic measurements showed that the experienced Germans produced the English /æ/ more accurately than did the inexperienced group. One of the major findings was that adult L2 learners could learn to produce the new L2 vowels in a similar way to native speakers. Bohn and Flege (1992) concluded that the experienced group produced the /ɛ/ and /æ/ contrast similarly to English natives, while the inexperienced group was unsuccessful in producing the /ɛ/ and /æ/ contrast.

Previous research has also examined the impact of language use on L1 attrition (Schmid & Köpke, 2007; de Leeuw et al., 2009, Mayr et al., 2012) and points to a possible relationship between exposure to one's L1 and the maintenance of an L1 accent. De Leeuw et al. (2010) attempted to discover whether L1 German speakers living in the Netherlands or English-speaking Canada were thought to have a 'foreign' accent in German as rated by native German speakers. Nineteen German monolinguals listened to five German monolinguals' accents, 34 L1 German speakers who lived in Canada and 23 L1 German speakers who were residing in the Netherlands at the time. On average, the bilingual German speakers had emigrated at age 27 and had been living in their adopted country for 37 years. Of those living in Canada and the Netherlands, nine and five of the speakers, respectively, were considered to be non-native German speakers, judging from their pronunciation. De Leeuw et al. (2010) considered that consistent and good quality contact with the native language has an impact on participants' propensity for developing a 'global foreign accent' in their L1 because it delayed L1 attrition. Notably, good quality contact seemed more effective at preventing L1 attrition than did the amount of time spent in a foreign country, or even the age at the time of arrival. Specifically, they found that L1 use in communicative contexts where code-switching was more likely to occur was less likely to be a predictor of an increase in L1 'foreign accent ratings'; the less an individual used 'code switching', the less likely they were to be judged as having a foreign accent in their native language. In addition, Mayr et al. (2012) studied monozygotic twins and found that L2 long-term experience affected the L1 vowels production.

Generally, L1 attrition seems to be most greatly affected by examples of language use where code-switching is less likely, for example, in the home. Because these studies focus not on one specific area of pronunciation, but on a global foreign accent (see Chapter Two, section 2.2.3.1, for more details), it is difficult to identify specific aspects of pronunciation, which are affected by language use. To make headway into providing a more detailed account of the role of language use in L2 acquisition and L1 attrition in the phonetic domain, this chapter is concerned with only one aspect of pronunciation: shared vowels. Therefore, while the current study is narrower than most, it offers a greater level of depth in its investigation of the effect of language use on L2 vowel learning and L1 attrition.

Role of Aptitude in L2 Acquisition and L1 Attrition

Does accurately producing unfamiliar or new sounds require any special talent? Might learners with ‘a good ear’ or phonetic talent be better at maintaining a native accent in their L1 (and possibly also be better at acquiring near-native production in the L2)? These questions need to be addressed because there is still a scarcity of research investigating the phonetic aptitude and/or talent factor. Few previous investigations of aptitude have focused on L2/foreign language learning (Tahat et al., 1981; Flege et al., 1995; Thompson, 1995). As I stated in the previous chapter, while previous studies have used specific aptitude tests to explore aptitude in L2 acquisition (e.g. Jilka et al., 2008), to the best of my knowledge, there is no study examining the role of aptitude on L1 attrition that uses specific aptitude tests to measure the participants’ levels of aptitude. Carroll (1993) described foreign language aptitude (or phonetic talent) as a natural propensity for mastering an L2. However, an accurate predictor of foreign language aptitude does not appear to exist for any particular population (Novoa et al., 1988; Schneiderman & Desmarais, 1988; Ross et al., 2002; Skehan, 2002; Dörnyei & Skehan, 2003).

Most previous studies have used indirect measures of aptitude to study the effect of talent on language acquisition and/or L1 attrition (Suter, 1976; Neufeld, 1979; Flege et al.,1995; Flege et al., 1999; Hopp & Schmid, 2013). Jilka et al. (2008), one of the few studies using an actual aptitude test, found that phonetic talent predicted performance in L2 spoken tasks. This is in line with Mennen’s (2004) proposal that some people are simply ‘outliers’ because they have the uncommon ability to attain native-like aptitude in both their L1 and L2. The present study uses specific aptitude tests to examine the effect of phonetic aptitude on shared vowels between Arabic and English in the participants’ L1 attrition and L2 acquisition.

3.2.4 Arabic and English Vowel Systems

Numerous studies have examined the American and British English vowel systems (Peterson & Barney, 1952; Chomsky & Halle, 1968; Hillenbrand et al., 1995; Deterding, 1997; Watt, 2002; Labov et al., 2006), which are both characterised as large vowel systems comprising simple vowels and diphthongs. However, there are fewer studies on the Arabic vowel system, which is a smaller vowel system that also contains simple vowels and diphthongs (Al-Ani, 1970; Mitleb, 1984; Al-Otaibi, 1988; Alghamdi, 1998; Khattab 2007; Khattab & Al-Tamimi, 2008; Alotaibi & Hussain, 2009; Mugair & Mahadi, 2014).

Vowel quality and vowel quantity (length) are the two phonetic parameters that describe vowels. Vowel quality measures the places of articulation, the tongue's position and the lips' shape. Vowel quantity is measured as the duration of the phonetic segment. In this chapter, I focus on the quality of the vowels of Arabic and English. As shown in Table 3.1. there is a clear difference between the number of vowels in Arabic and English. While short and long vowels in English typically have distinct vowel qualities and can be distinguished in terms of vowel height (Kopczynsk & Meliani, 1993), length is the distinctive feature for short and long vowels in Arabic (Kopczynsk & Meliani, 1993). In the case of high front vowels in Arabic though, the /ɪ/ and /iː/ do not only differ in duration, but also in vowel quality.

It should be mentioned that previous studies on English and Arabic seem to use slightly different IPA symbols for similar vowels. For example, previous studies on Arabic use /aw/, as in /jawm/ (Engl.: *day*), whereas previous studies on English use /əʊ/, as in /həʊs/, for quite similar diphthongs (Brierley et al., 2016; Mayer et al., 2012). The current study uses the IPA symbols that are used in the previous literature on both languages, so that vowels that are similar across both languages may nevertheless be represented by different IPA symbols here.

It should also be noted that the focus of the current chapter is on vowel quality, not vowel quantity. Including vowel quantity would have been beyond the scope of the current chapter. However, the vowel data are fully annotated so that duration information can be easily extracted and can be analysed in a future study.

Table 3.1. Vowel Inventories of MSA and SSBE with Vowels Represented in IPA

Vowels	***Arabic***	***English***
Short	ɪ, a, u	ɪ, e, æ, ə, ʌ, ʊ, ɒ
Long	iː, aː, uː	iː, ɜː, ɑː, uː, ɔː
Diphthong	aw, aj	eɪ, aɪ, ɔɪ, əʊ, ɑʊ, ɪə, eə, ʊə

Arabic is a Semitic language, characterised by a rich system of consonants but a limited vocalic system. The most common term used to stand for the Quran Classical Arabic, which is known in Arabic as /fusˀða/ (pure speech), is MSA. MSA has three short vowels (/a/, /ɪ/ and /u/) called /ħarakaːt/ حركات and 3 long vowels (/aː/, /iː/ and /uː/; Shah & Shah, 2007) called /uruːf madd/ حروف مدّ. The length of the vowel differentiates word meaning; for example, /dʒamal/ جَمَل "camel" and / dʒamaːl/ جمال "beauty". Arabic has two diphthongs, (/aj/ and /aw/), which occur in non-final positions (Prochazka, 1988; Alotaibi & Husain, 2009); examples of these are /jawm/ "day" and /bajt/ "house". As the English language has more vowels than Arabic, I focus only on long and short vowels that are shared between Arabic (/ɪ/, /a/, /u/, /iː/, /aː/ and /uː/) and English (/a/, /ɪ/, /ʊ/, /ɑː/, /iː/ and /uː/) in this chapter. Table 3.2 lists the vowels that are shared between Arabic and English and provides example words in both languages.

Table 3.2. Arabic and English Shared Vowels with Examples

Arabic	***English***
(1) */a/ low front short as in /rad/ "reply"* رد	(1) /a/ low front short as in had
(2) */aː/ low front long as in /laː/ "no"* لا	(2) /ɑː/ low front long as in hard
(3) */ɪ/ high front short as in /sɪn/ "tooth"* سِن	(3) /ɪ/ high front short as in kick
(4) */iː/high front long as in /fiːl/ "elephant"* فيل	(4) /iː/ high front long as in feet
(5) */u/ high back short as in /hum/ "they"* هم	(5) /ʊ/ high back short as in should
(6) /uː/high back long as in /nuːn/ letter "n" نُون	(6) /uː/ high back long as in cool

Compared to other languages such as English and German, Arabic is less researched. Al-Ani's (1970) research was the first acoustic study of Arabic vowels, and Ghazeli (1979), Belkaid (1984), Mitleb (1984) and Lahlou (1981-1982) subsequently examined vowels in Arabic. These studies all involved MSA, using speakers from various dialect areas (see Newman & Verhoeven, 2002, for more information and a comparative list of vowel frequencies). In general, the length of the sound (phonemic length) has been determined to have a different role in the Arabic language (Mitchell, 1993; Newman & Verhoeven, 2002). Khattab (2007) and Khattab and Al-Tamimi (2008) studied the temporal opposition of Arabic long vs short vowels and consonants, and they suggested that long vowels are about twice as long as short vowels in Lebanese Arabic. Similar to English (Wells, 1962; Deterding, 1997, 2006), vowel duration is affected by the voicing of the following consonant in certain varieties of Arabic (Almbark, 2012), such that vowels are longer before voiced consonants than before voiceless consonants (Denes, 1955; Mitleb, 1982; Munro, 1993).

Table 3.3 shows the F1 and F2 formant values for the Arabic vowels /iː/, /ɪ/, /uː/, /u/, /aː/ and /a/ as measured in various previous studies (adopted from Newman, 2005, and Alotaibi & Hussain, 2009).

Table 3.3. Mean Formant Frequencies of the Six Phonetic Short and Long Vowels of a Comparison of Arabic

Reference	**Dialect**	**/iː/**		**/ɪ/**		**/uː/**		**/u/**		**/aː/**		**/a/**	
		F1	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
Al-Ani (1970)	Iraqi	285	2200	290	**2200**	285	775	290	800	675	1200	600	1500
Ghazeli (1979)	Algeria, Tunisia, Libya, Egypt, Jordan and Iraq	310	2225	455	1780	330	900	450	1125	-	-	-	-
Belkaid (1984)	Tunisian	285	2195	355	1830	310	790	340	995	425	1720	400	**1640**
Haidar (1994)	Qatar, Lebanon, Saudi Arabia, Tunisia, Syria, Sudan, United Arab Emirates and Jordan	315	2230	485	1750	355	835	**500**	1120	**690**	**1500**	**675**	1585

Newman & Verhoeven (2002)	Quraic and Cairene / one male for each dialect	390	1725	440	1770	**470**	**1120**	480	**1170**	620	1455	616	1560
Alotaibi & Hussain (2009)	MSA, 10 male (nine from different regions in Saudi Arabia and one from Egypt) and one child	**412**	**2132**	479	1545	429	859	489	975	684	1193	591	1102

Note: Adopted from Newman (2005) and Alotaibi & Hussain (2009). Values in bold are the highest in the range and empty spaces indicate absence of data.

Wells (1962), Deterding (1997, 2006), and Hawkings and Midgley (2005) have measured vowel formants for English. Table 3.4 and Table 3.5 show the F1 and F2 values from these studies for the six vowels that are relevant to the current study.

Wells (1962) examined formant values of British English vowels for 25 male native speakers of English (British) using the words *heed, hid, head, had, hard, hod, haw'd, hood, who'd, hud* and *heard*. Similarly, Deterding (1997) measured formants of SSBE vowels produced by five female and five male BBC broadcasters. The results from the experiment showed that male's vowel pronunciations were significantly less peripheral in connected speech than in the measurements taken from the citation words.

Table 3.4. Mean Formant Frequencies of the Six Phonetic Short and Long Vowels of a Comparison of English, as Produced in British and American English

Reference			/iː/		/ɪ/		/uː/		/ ʊ /		/ɑː/		/a/	
	Dialect	**Gender**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
Deterding (2006)	British	M	280	2249	367	1757	316	1191	379	1173	646	1155	690	1550
		F	303	2654	384	2174	328	1437	410	1340	910	1316	1018	1799
Wells (1962)	British	M	285	2373	356	2098	309	939	376	950	677	1083	748	1746

Note: Adopted from Deterding (2006).

Table 3.5. *Mean formant frequencies per age group adopted from Hawkings & Midgley (2005).*

Vowel	**65+**		**50–55**		**35–40**	
	F1	**F2**	**F1**	**F2**	**F1**	**F2**
/iː/	285	2283	269	2355	269	2312
/ɪ/	382	2024	341	2074	374	2115
/ɑː/	665	1085	639	1041	608	1062
/a/	518	875	522	865	496	833
/ʊ/	376	990	371	975	381	984
/uː/	301	994	283	1112	288	1336

To align with the current study's purpose, I focus on the British English and the MSA whose regional dialects Saudi Arabic. Hence, I compared the participants from the study conducted by Alotaibi and Hussain (2009) with the participants from Deterding's (2006) study. The values given in Tables 3.3 and 3.4 suggest that the vowels /iː/, /ɪ/, /uː/ and /u/ in Arabic have a higher F1 value than they do in English, which means that the Arabic vowels /iː/, /ɪ/, /uː/ and /u/ are produced using a lower tongue position than is required by their English counterparts. By contrast, the Arabic vowels /aː/ and /a/ have lower F1 values than do the English vowels, which means that the Arabic vowels /a/ and /aː/ are produced using a higher tongue position than /a/ and /ɑː/ in English. Thus, overall, the English vowels have more extreme positions in vowel space in terms of vowel height, possibly because English has more vowels than Arabic and needs to use more of the vowel space to distinguish vowels from adjacent vowels. In addition, the values shown in Tables 3.3 and 3.4 suggest that the vowels /iː/, /ɪ/, /uː/, /u/, /aː/ and /a/ in Arabic have lower F2 values compared to their English counterparts, which means that these English vowels are produced using a more frontal tongue position than when produced in Arabic. To summarise, even though both English and Arabic have the vowels /iː/, /ɪ/, /uː/, /ʊ/or /u/, /ɑː/ or /aː/ and /a/, the F1 and F2 values of these vowels differ between the two languages.

3.3 The Present Study

The first aim of the present chapter is to investigate whether highly-fluent late consecutive A-E bilinguals who moved to the UK in adulthood and highly-fluent late consecutive E-A bilinguals who moved to the KSA or Yemen in adulthood show native-like vowel productions in their L2 and/or show L1 attrition of vowels. In the present chapter, the production of vowels is the phonetic variable, selected because the literature indicates differences between English and Arabic in the vowel's phonetic qualities. Furthermore, few studies have been conducted on the production of vowels in Arabic. Thus, the present chapter attempts to narrow this gap in the literature. This chapter also attempts to test whether conclusions drawn from previous studies hold true across two typologically different language groups such as Arabic and English.

The second aim is to determine whether language use, as measured by a detailed questionnaire, has an effect on L1 attrition and the L2 acquisition of vowels. Many studies assume that the prolonged use and exposure to L2 leads to L1 attrition; thus, language use was selected as one of the main factors in this investigation. All participants in the current study are late L2 learners who have lived in the L2 environment for a substantial amount of time and thus have extensive exposure to the L2.

The final aim of this chapter is to ascertain whether phonetic aptitude plays a role in L2 language acquisition and L1 attrition of vowels. Following Jilka (2007, 2009) and Jilka et al. (2008, 2011), aptitude was measured through a variety of phonetic aptitude tests. As mentioned in Chapter Two, participants with high phonetic aptitude may show more native-like pronunciation in their L2, but this may happen to the detriment of pronunciation in the L1. However, learners who exhibit a specific aptitude for the phonetic aspects of language may resist L1 attrition better than others with less of a phonetic aptitude in the L2. That is, those learners with 'a good ear' may be better at maintaining a native accent in their L1 while also achieving more native-like L2 productions (cf. Mennen, 2004).

3.4 Hypotheses

1. Arabic and English: Based on the previous literature (e.g. Deterding, 1997, Alotaibi & Hussain, 2009), I would expect that

a. Arabic native speakers tend to produce /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position (lower Z3-Z1) than that used in English.
b. Arabic native speakers tend to produce /ɑː/ and /a/ with a higher tongue position (higher Z3-Z1) than that used in English.
c. Arabic native speakers tend to produce /ɪ/, /iː/ /uː/, /ʊ/, /a/ and /ɑː/ with a less frontal tongue position (higher Z3-Z2) than that used in English.

2. L2 acquisition: The SLM predicts that participants would not form a new vowel category for L2 vowels that are similar to their L1 vowels. I therefore expect the bilingual participants will to pronounce the L2 vowels as they do in their L1 and that the bilinguals' L2 vowels are expected to differ from those of the monolingual participants.

3. L1 attrition: Based on the previous literature and assumptions from the SLM, I would expect that the A-E and E-A bilinguals have merged vowel categories for the similar vowels considered in this chapter. If this merged category reflects the L1, then participants would not show attrition in their L1. However, if this merged category has shifted towards the L2 due to prolonged L2 exposure, then I would expect to see evidence of L1 attrition and possibly values that are typical for neither the L1 nor the L2.

4. Language use: Based on the existing literature, I expect an influence of language use on L2 acquisition and L1 attrition, such that bilinguals with more L2 use will produce more native-like vowels in their L2 compared to those with less L2 use. Conversely, more L1 use will be better suited to inhibit the effects of L1 loss.

5. Phonetic aptitude: While there is no consensus in the literature regarding the role of phonetic aptitude in L2 acquisition and L1 attrition, I tentatively assume that participants with a high phonetic aptitude will show more native-like production of vowels in both their L1 and L2 (cf. Mennen, 2004).

3.5 Methods

3.5.1 Participants

The participants for this chapter were the same as those used in Chapter Two. One additional participant, who had to be excluded for the analyses in Chapter Two, is included here. Table 3.6 summarises the participant characteristics.

Table 3.6. Summary of Participant Characteristics

Participant group	N (gender)	Mean age (SD)	Mean AOA (SD)	Mean LOR (SD)
Monolingual Saudi, native Arabic speakers	15 (4 males, 11 females)	36.06 (SD = 6.91)	NA	NA
Monolingual native English speakers	15 (2 males, 13 females)	40.93 (SD = 9.65)	NA	NA
A-E bilingual	15 (one male, 14 females)	39.4 (SD = 3.75)	19 (SD = 2.8)	20 years (SD = 2.6)
E-A bilingual	15 (four males, 11 females)	33.66 (SD = 3.9)	16.7 (SD = 1.12)	17 years (SD = 3.9)

3.5.2 Materials and Procedure

Production of Vowels

The first three formants (F1, F2 and F3) for the English and Arabic vowels /iː/, /ɪ/, /uː/, /ʊ/, /ɑː/ and /a/ were used for the analyses in this chapter. I have normalized vowels with the Bark Difference Method (Syrdal & Gopal, 1986), where vowel height is modelled as Z3-Z1 (Bark-converted F3 minus Bark-converted F1) and vowel frontedness is modelled as Z3-Z2 (Bark-converted F3 minus Bark-converted F2). These values relate differently to tongue height and frontedness than the raw F1 and F2 values. Specifically, the higher Z3-Z1, the higher the tongue and the higher Z3-Z2, the more backed the vowel. The Bark Difference method was chosen here because it is rather common and because it is a vowel-intrinsic method and can therefore be used on a subset of vowels that make up the vowel inventory of a given language or languages (Clopper, 2009). Speech production data of Arabic and English vowels were collected from the participants. Vowels were embedded in the /hVd/ frame to control for the effect of phonetic context. The use of the /hVd/ frame meant that participants produced both words and non-words. Table 3.7 shows the Arabic and English target words of the shared vowels (short and long) and their phonetic symbols. Two of the Arabic words are not real words in Arabic. The participants produced the Arabic words (*hɪd**, *h*ad, *hud**, *hiːd*, *haːd* and *huːd*; where * indicates a non-word) and English words (*hɪd*, *h*ad, *h*ʊ*d*, *hiːd*, *h*ɑː*d* and *huːd*) both in isolation and in a carrier phrase (see Appendix D). The participants read each word six times in isolation and in the carrier phrase. To ensure activation of the intended categories, especially in the case of non-words, the target words (see Table 3.2) were primed with real words that had the same vowel sound, as in examples (3.1) to (3.4):

(3.1) Arabic in isolation:
*qud, sˤud, hud**
drive, block, hud*

(3.2) Arabic in a carrier phrase:
qultʊ qud θuma ðahbt. qultu sˤud θuma ðahbt. qultu hud θuma ðahbt.
"I said drive, then I left. I said block, then I left. I said hud*, then I left."

(3.3) English in isolation:
would, could, hood

(3.4) English in carrier phrase:
I said would and then I left. I said could and then I left. I said hood and then I left.

A native speaker of each language checked the words and sentences to make sure that they sounded as natural as possible and contained no mistakes.

Table 3.7. Target Words and Phonetic Symbols (indicates non-words)*

Vowels	Arabic (MSA)	English (SSBE)
Short vowels	ɪ /hɪd/* "هِد"	ɪ /hɪd/ "hid"
	a /had/ demolish "هَد"	a / had/ "had"
	u /hud/* "هُد"	ʊ /hʊd/ "hood"
Long vowels	iː / hiːd/* "هيد"	iː /hiːd/ "heed"
	aː /haːd/ guidance "هاد"	ɑː /hɑːd/ "hard"
	uː /huːd/ a name "هود"	uː /huːd/ "who'd"

Background Questionnaire

I used information from the language background questionnaire described in Chapter Two. As already mentioned in section 2.5.2., the purpose of the questionnaire was to collect sociolinguistic background information from the bilingual participants. The questionnaire included demographic questions, and the bilinguals answered questions about contact times with speakers of each language, the frequency and type of use of each language, and their attitudes towards the relevant languages and cultures. See Chapter Two for a more detailed description of the questionnaire.

Language Proficiency and Aptitude

I also used the results from the proficiency and aptitude tests described in Chapter Two. As mentioned in Chapter Two, bilinguals' Arabic and English language proficiency was measured using the TOEFL (computer-based version) and an APT. The tests were limited to a grammar (structure) test and a listening comprehension test; both tests require vocabulary knowledge and grammatical judgment. For more details, see Chapter Two, section 2.5.2.

The aptitude test was adapted from MLAT and PLAB samples, which are available online. This chapter used the MLAT sample questions in Part III (Spelling Cues, which tests sound-symbol associations), Part V (Paired Associates, which tests memory for novel words) and the PLAB sample questions in Part V (Sound Discrimination, which tests discrimination of pitch, orality and nasality). For a more detailed description of the aptitude test, see Chapter Two, section 2.5.2.

3.5.3 Overall Procedure

As mentioned in Chapter Two, the data collection was carried out over six months in KSA and the UK and was mainly done through the snowball method. Participants were emailed information about the study and a consent form. After giving their consent, the bilingual participants attended three sessions on separate days up to one week apart. The first session comprised the L2 proficiency assessment (in the L2) followed by the APT. The second and third sessions comprised three different production tasks, with Arabic productions being recorded in one session and the English productions in the other. All productions were recorded with a handheld Sony tape recorder in a quiet environment. During these sessions, the participants read a list of words and sentences containing different vowels. The participants first read the lists of words, such as *would, could* and *hood*, six times, and then read these words embedded in sentences six times, resulting in 36 vowel tokens ($6 \times 6 = 36$) in isolation and 36 ($6 \times 6 = 36$) tokens in a carrier phrase for each participant and each language.

The participants then read the question-answer dialogs analysed in Chapter Two and a short cartoon, which allowed for the production of plosives and which are analysed in Chapter Four. The monolingual participants were required to attend only two sessions because they only participated in the production tasks in their native language.

3.5.4 Data Analysis

To analyse the vowels in the target words in the carrier phrases and in isolation, I first determined the beginnings and ends of vowels using Praat and then labelled the vowels; both processes were done manually. Then a script in Praat extracted duration and formant information (Boersma & Weenink, 2016). The vowel duration was measured from the approximate beginning of the selected vowel (the first peak in the digitised waveform) to the end of the vowel. Next, a Praat script extracted F1 and F2 for all the selected vowels, spoken in isolation and in the carrier phrase. The automatically extracted formants were manually checked for possible tracking errors (such as halving and doubling) and were corrected if needed.

Linear mixed effects models (Bates et al., 2015) were used to analyse vowel mid-point values. Linear mixed effects models allow both fixed and random effects to be included in the analysis. The analyses were done using RStudio (Version 1.1. 456, 2018) and the lme4 package (Bates et al., 2015) (Version 1.1-18-1, 2018). I analysed vowel midpoint values separately for each of the six vowels (/ɪ/, /iː/, /a/, /ɑː/ or /aː/, /ʊ/ or /u/ and /uː/). Recall that the literature in Arabic and English uses different IPA symbols for some of the vowels, so for easier comparison I will be using only the English IPA symbols in the analysis. Depending on the analysis, the fixed effects in the analyses included language (Arabic vs English), group (Arabic monolinguals (A mono), English monolinguals (E mono), A-E and E-A bilinguals), speaking condition (in isolation vs in a carrier phrase), language use (more Arabic vs more English) and aptitude (high vs low). Participant and item were added as random effects. F1 and F2 values were converted into Z3-Z1 and Z3-Z2 respectively.

The bilingual participants for this experiment were the same as those used in Chapter Two. Table 3.8 summarises the proficiency, aptitude and language use among the A-E and E-A bilinguals.

Table 3.8. Proficiency, Aptitude and Language Use Among A-E and E-A Bilinguals

Measure	**A-E bilinguals**	**E-A bilinguals**	***t*-test *p*-value**
Proficiency	TOEFL: mean = 42.1 (SD = 1.9; range = 39-45)	APT: mean = 43.7 (SD = 1.0; range = 42-46)	NA
Aptitude spelling cues	mean = 17.5 (SD = 4.2; range = 11-25)	mean = 19.8 (SD = 3.3; range = 14-24)	t = -1.4494 p = .16
Aptitude paired associates	mean = 14.6 (SD = 2.0; range = 12-18)	mean = 15.6 (SD = 2.2; range = 12-19)	t = -1.4155 p = .17
Aptitude sound discrimination	mean = 23.5 (SD = 2.3; range = 20-28)	mean = 23.5 (SD = 2.7; range = 20-28)	t = 0.1883 p = .85
Aptitude total	mean = 55.6 (SD = 7.8; range = 45-71)	mean = 58.9 (SD = 7.5; range = 49-69)	t = -1.0437 p = .31
Language use	More Arabic = 8 More English = 7	More Arabic = 5 More English = 10	NA

3.6 Results

Overview

Before presenting the statistical results, I provide an overview of the Arabic and English vowels' mean F1 and F2 values for all participant groups (see Tables 3.9 and 3.10). F1and F2 values are presented here for easier comparison with the F1 and F2 values reported in previous studies.

Table 3.9. Mean and Standard Deviation for F1 and F2 for Each Arabic Vowel Produced by A Mono, A-E Biling and E-A Biling

Vowel		**/a/**		**/aː/**		**/ɪ/**		**/iː/**		**/u/**		**/uː/**	
Formants		**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
A mono	**Mean**	669	1645	753	1566	573	1919	559	2194	572	1419	540	1200
	SD	164	304	184	265	230	412	319	570	194	243	211	431
A-E biling	**Mean**	711	1643	804	1635	488	2102	383	2438	486	1311	419	56.3

	SD	104	363	135	233	57.8	437	57.9	651	58	148	972	206
E-A biling	**Mean**	706	1689	741	1681	489	2129	401	2466	495	1294	427	1011
	SD	79.8	321	103	307	73	317	57.9	444	62.8	219	54.4	316

Table 3.10. Mean and Standard Deviation for F1 and F2 for Each English Vowel Produced by E Mono, A-E Biling and E-A Biling

Vowel		**/a/**		**/ɑː/**		**/ɪ/**		**/iː/**		**/ʊ/**		**/uː/**	
Formants		**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
E mono	**Mean**	797	1531	780	1280	523	2148	374	2475	530	1419	393	1668
	SD	131	203	116	219	60.5	382	61.3	490	96.4	381	53.1	494
A-E biling	**Mean**	860	1650	804	1288	506	2077	375	2506	494	1350	407	1164
	SD	174	300	104	277	71.5	409	59.8	614	80.7	301	55.9	245
E-A biling	**Mean**	808	1592	753	1249	514	2122	378	2409	504	1373	400	1309
	SD	144	206	98.4	128	74.9	310	64	608	87.6	227	61.5	386

The values in Tables 3.9 and 3.10 show that the vowels /ɪ/, /iː/ and /uː/ in Arabic have numerically lower F2 values and are thus less fronted than their English counterparts. By contrast, /a/ and /ɑː/ have numerically higher F2 values and are thus more fronted in Arabic compared to their counterparts in English. Both monolingual groups produced the /ʊ/ with the same mean F2 value. Additionally, the F2 values for /ɪ/, /ɑː/ and /uː/ are similar to Deterding's (2006) and Alotiabi and Hussain's (2009) studies, which show that F2 values in Arabic are numerically lower than they are in English, while the F2 values for /a/ and /ʊ/ reveal a pattern in this study that is different to these two previous studies. Specifically, the current study shows that the Arabic /a/ has a numerically higher F2 value than in English, which contrasts with both Deterding's (2006) and Alotiabi and Hussain's (2009) results. Further, the present study shows that /ʊ/ has the same F2 value in both English and Arabic, while in Deterding's (2006) and Alotiabi and Hussain's studies, the F2 value of /ʊ/ is higher in English than it is in Arabic.

I now consider the bilingual participants, starting with the A-E bilinguals and the Arabic vowels. The values in Table 3.9 show that the A-E bilinguals produced the vowels /a/ and /ɑː/ in Arabic with a numerically higher F1 and thus lower tongue position than did the Arabic monolinguals, and they produced the /ɪ/, /iː/, /ʊ/ and /uː/ with a numerically lower F1 and thus higher tongue position than did the Arabic monolinguals. Like the English monolinguals, the A-E bilinguals are using a slightly larger vowel space in terms of tongue height than are the Arabic monolinguals. Both A monolinguals and A-E bilinguals produced the /a/ and /ɑː/ with almost the same mean F2, whereas the A-E bilinguals produced the /ɪ/, /ʊ/ and /u/ with a numerically lower F2 (less fronted tongue position). The F2 of Arabic /iː/ produced by the A-E bilinguals is numerically higher than that produced by the A monolinguals, suggesting a more fronted tongue position.

Next, I consider the E-A bilinguals and the Arabic vowels. Table 3.9 shows that the E-A bilinguals produced the Arabic /a/ with numerically higher F1 and thus lower tongue position, and they produced the Arabic /ɑː/, /ɪ/, /ʊ/ and /uː/ with a numerically lower F1 and thus a higher tongue position compared to the A monolinguals. Similar to the E monolinguals and the A-E bilinguals, the E-A bilinguals are using an overall slightly larger vowel space in terms of tongue height than are the A monolinguals. With regards to the F2, E-A bilinguals produced the Arabic /a/, /ɑː/, /ɪ/ and /iː/ with numerically higher F2 values than the A monolinguals, meaning they produced them with a more fronted tongue position. By contrast, E-A bilinguals produced the /ʊ/ and /uː/ with numerically lower F2 values and thus less fronting than the A monolinguals. Thus, the E-A bilinguals use a slightly larger vowel space in terms of frontedness and backness than do the Arabic monolinguals.

Moving on to the A-E bilinguals and the English vowels, Table 3.10 illustrates that the A-E bilinguals produced the English /a/, /ɑː/ and /uː/ with numerically slightly higher F1 values and thus a lower tongue position than did the E monolinguals and they produced the /i/ with almost the same mean F1 value as the E monolinguals. However, the A-E bilinguals produced the English /ɪ/ and /ʊ/ with numerically lower F1 values and thus a higher tongue position than did the E monolinguals. In addition, the A-E bilinguals produced the /ɪ/, /uː/ and /ʊ/ with numerically lower F2 values and thus a less fronted tongue position and the /a/, /ɑː/ and /iː/ with numerically higher F2 values and thus a more fronted tongue position than did the E monolinguals.

Finally, I consider the E-A bilinguals and English vowels. Table 3.10 shows that the E-A bilinguals produced /a/, /iː/ and /uː/ with numerically slightly higher F1 values and thus a slightly lower tongue position than did the E monolinguals. However, the E-A bilinguals produced the /ɑː/, /ɪ/ and /ʊ/ with numerically lower F1 values and thus a higher tongue position than did the English monolinguals. Interestingly, these findings suggest that the E-A bilinguals use a slightly larger vowel space in terms of tongue height than do the English monolinguals for the short/lax vowels, but a slightly smaller vowel space in terms of tongue height for the long/tense vowels. The E-A bilinguals produced only /a/ with numerically higher F2 and thus a more fronted tongue position than did the English monolinguals, whereas the F2 for the rest of the vowels, i.e. /ɑː/, /ɪ/, /iː/, /ʊ/ and /uː/, was numerically lower, indicating less frontedness than for the E monolinguals.

I now continue with the statistical analysis. Unlike the information shown in Tables 3.9 and 3.10, which mainly serves as a comparison with previous studies reporting F1 and F2 values, the statistical analysis used the normalized Z3-Z1 and Z3-Z2 values. Specifically, I present the statistical results of the comparison between the A monolinguals (A mono), English monolinguals (E mono) and the A-E and E-A bilinguals in producing shared vowels (/a/, /ɪ/, /ʊ/, /ɑː/, /iː/ and /uː/). First, I compared the monolingual patterns for each vowel separately to determine whether the Z3-Z1 and Z3-Z2 (Bark) of the monolinguals differ across the two languages. If they differed, I compared the bilingual and monolingual productions in Arabic and English to determine whether the bilinguals differed from the monolingual groups. I examined the normalised first and second formants (Z3-Z1 and Z3-Z2 Bark) of the short vowels (/a/, /ɪ/ and /ʊ/), first in Arabic and then in English. I then examined the formant values for the long vowels (/ɑː/, /iː/ and /uː/), first in Arabic and then in English. Finally, I tested whether significant differences in vowel production between bilinguals and monolinguals could be attributed to language use and phonetic aptitude.

I will refer to differences both in terms of Bark and of tongue height and frontness/backness, even though the data is acoustic rather than articulatory and there is no one-to-one relationship between the Bark values and the tongue height and frontness/backness.

3.6.1 Z3-Z1 and Z3-Z2 (Bark) for the /a/ Vowel in Arabic and English Across Speaker Groups

I begin my analysis by considering the short vowel /a/. Figures 3.1 and 3.2 show the mean Z3-Z1 and Z3-Z2 (Bark) for the /a/ vowel produced by A mono, A-E bilinguals and E-A bilinguals. The normalised first formant (Z3-Z1) represents the vertical dimension and the normalised second formant (Z3-Z2) represents the horizontal dimension. Close examination of Figures 3.1 and 3.2 shows that the vowel /a/ differs in Arabic and English, especially with respect to Z3-Z1. The A-E and E-A bilinguals' /a/ production in Arabic conforms closely to the native norm for the Arabic /a/ vowel. In addition, the production of the English /a/ by the A-E and E-A bilinguals conforms closely to the English native norm. However, A-E bilinguals have a greater range of tongue positions when producing the English /a/ vowel than do native English speakers.

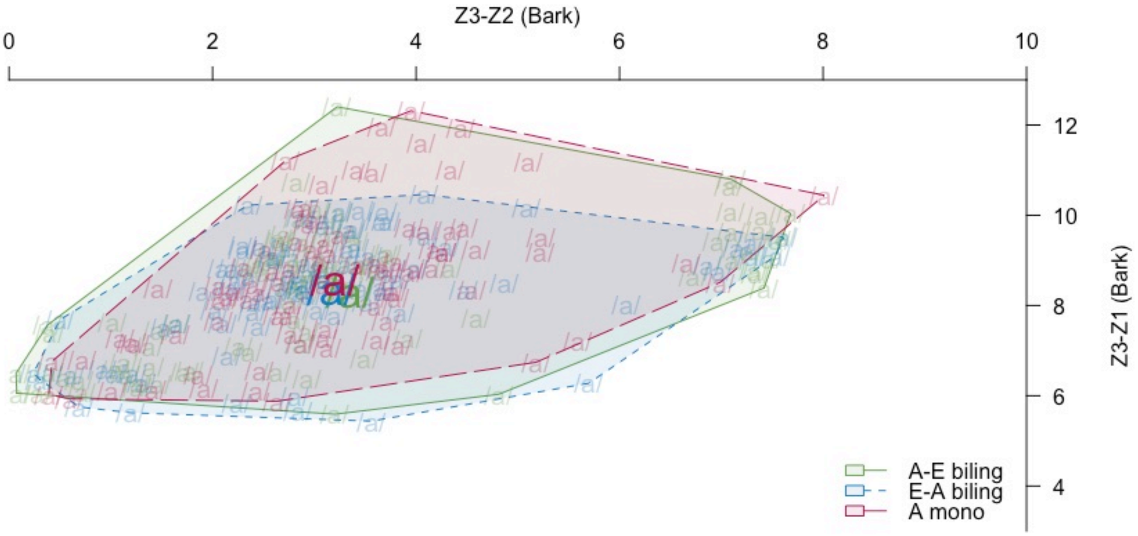


Figure 3.1. Mean Z3-Z1 and Z3-Z2 (Bark) for the Arabic /a/ vowel produced by A mono, A-E bilinguals and E-A bilinguals.

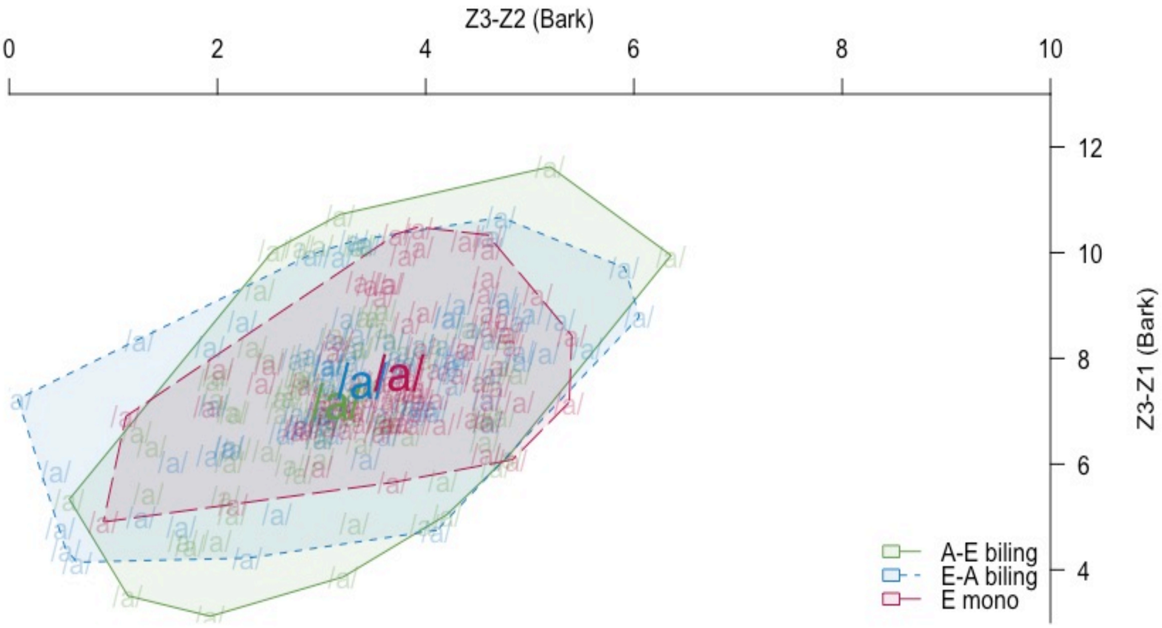


Figure 3.2. Mean Z3-Z1 and Z3-Z2 (Bark) for the English /a/ vowel produced by E mono, A-E bilinguals and E-A bilinguals

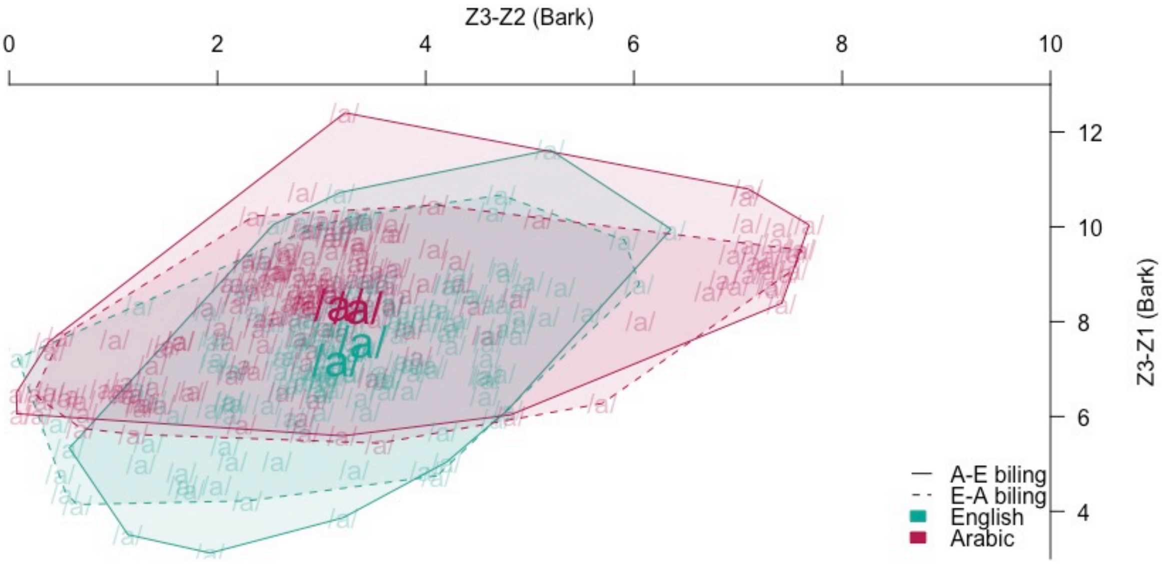


Figure 3.3. Mean Z3-Z1 and Z3-Z2 (Bark) for the /a/ vowel produced by A-E bilinguals and E-A bilinguals in Arabic and English

Results for the Z3-Z1 Values of the /a/ Vowel in Arabic and English

Comparison of Monolingual Patterns

First, I determined whether A mono and E mono differed in their productions of the /a/ vowel in terms of Z3-Z1 (tongue height). Figure 3.3 shows the Z3-Z1 values (associated with tongue height) of the Arabic and English /a/ as produced by A mono, E mono, A-E bilinguals and E-A bilinguals. I first compared the Z3-Z1 for A mono and E mono to determine whether tongue height differed between the two monolingual participant groups for the /a/ sound. Figure 3.4 shows that the Z3-Z1 for E mono was numerically lower than that for A mono speakers, meaning that E mono tend to produce the /a/ vowel with the tongue in a lower position than do A mono.

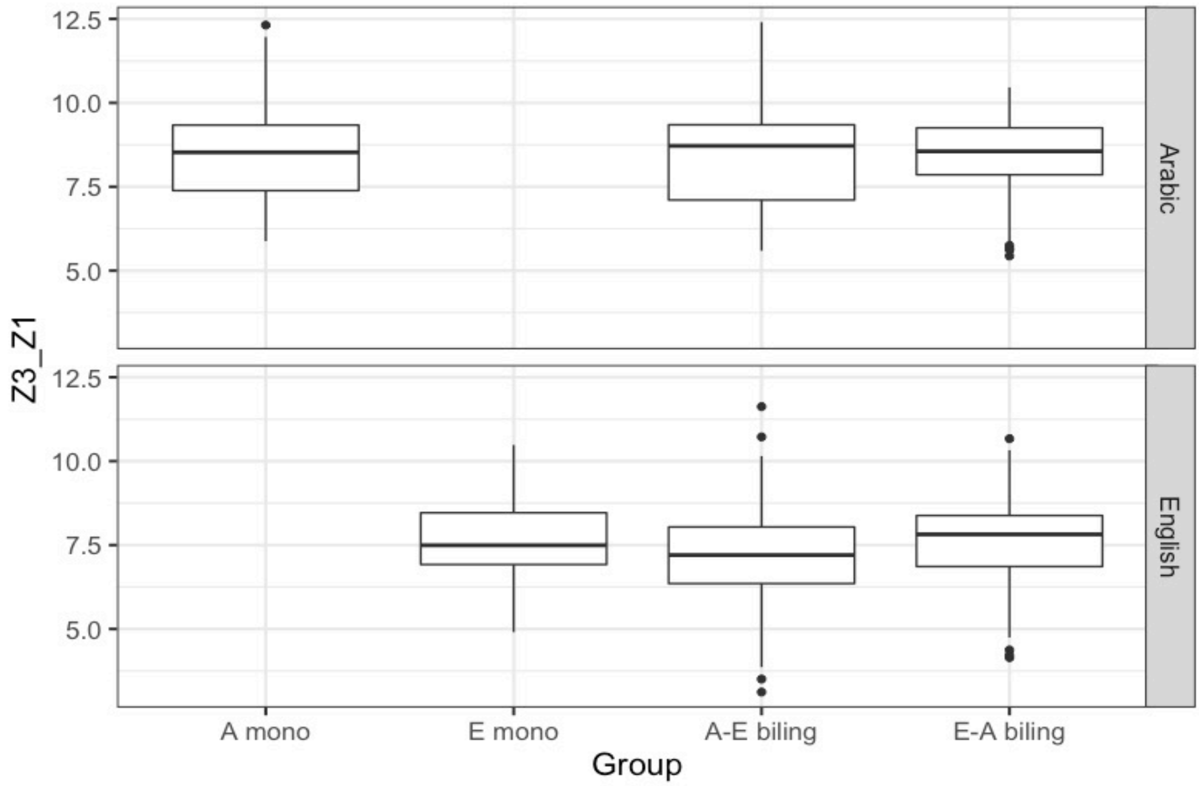


Figure 3.4. Boxplots for the Z3-Z1 values (Bark) of the Arabic and English /a/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

Linear mixed effects models were used to determine whether (1) A mono and E mono differed statistically significantly in terms of tongue height when producing the /a/ vowel and whether (2) the vowel would differ in isolation compared to when it is produced in a carrier phrase. The response variable for the analysis was the Z3-Z1. The fixed effects were participant group (A mono and E mono), speaking condition (in isolation vs in a sentence) and the participant group by speaking condition interaction. Table 3.11 shows the results. The analysis revealed that speaker group had a significant effect, indicating that the tongue height was significantly lower for E mono than for A mono speakers. Moreover, speaking condition had a significant effect, revealing that speakers had a significantly higher tongue height in the sentence-speaking condition than in the isolation speaking condition when producing /a/. The speaker group by speaking condition interaction was not significant.

Table 3.11. Z3-Z1 Comparison between A Mono and E Mono Producing /a/

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	8.404	0.2450	41.963	34.307	< .001
Speaker group	0.930	0.257	0.2578	-3.610	< .001
Speaking condition	0.5236	0.2458	154.220	2.130	= .035
Speaker group x speaking condition	-0.251	0.3502	153.862	-0.717	= .474

Bilinguals

Comparison of Z3-Z1 in E-A and A-E Bilinguals with A Monolingual Producing /a/ in Arabic

Since Arabic and English monolingual participants differed in producing the /a/ vowel, I compared the bilinguals' Arabic productions with the Arabic monolinguals' productions. Figure 3.4 above shows similar Z3-Z1 values of A mono, E-A bilinguals and A-E bilinguals when producing the Arabic /a/ vowel. Linear mixed effects models were employed to determine whether the E-A and A-E bilinguals shown in Figure 3.4 differ statistically significantly in terms of tongue height for the Arabic /a/. In addition, the model tested whether the vowels produced in isolation differ when produced in a carrier phrase. Treatment coding was employed to compare each bilingual group with monolinguals directly. The results in Table 3.12 revealed that no bilingual group differed from the A mono speakers in terms of tongue height for the Arabic /a/. Both bilingual groups had target-like productions in terms of the Z3-Z1 in Arabic. Speaking condition had no effect, revealing that the participants produced the /a/ vowel with similarly tongue height in the sentence-speaking condition and in the isolation speaking condition.

Table 3.12. Z3-Z1 Comparison between A Mono and E-A and A-E Bilinguals in Producing the Arabic /a/

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	8.336	0.239	51.168	34.767	< .0001
A mono vs A-E biling	-0.12619	0.265	248.538	-0.475	= .635
A mono vs E-A biling	-0.217 -	0.263	248.365	0.826	= .409
Speaking condition	0.504	0.260	248.298	1.935	= .054

A mono vs A-E biling x speaking condition	0.317	0.370	248.989	0.858	= .391
A mono vs E-A biling x speaking condition	0.031	0.371	248.351	0.100	= .920

Comparison of Z3-Z1 in A-E and E-A Bilinguals with E Monolingual Producing /a/ in English

As the production of the /a/ vowel differed between monolingual speakers, I compared the bilinguals' English productions with those of the E mono speakers. Figure 3.4 above displays similar Z3-Z1 values among the E mono and A-E and E-A bilinguals when producing the English /a/ vowel. Again, I used linear mixed effects models to determine whether the monolingual English and E-A and A-E bilinguals shown in Figure 3.4 differ statistically significantly in terms of tongue height for the English /a/ and to determine whether the vowels produced in isolation differ from those produced in a carrier phrase. The results show that E-A and A-E bilinguals do not differ from E mono in terms of tongue height. Additionally, the speaking condition did not affect Z3-Z1 for the English productions. Hence, both bilingual groups had target-like productions in terms of Z3-Z1 in English. The speaking condition had no effect on the bilingual groups or E mono speakers.

Table 3.13. Z3-Z1 Comparison between E Mono and A-E and E-A Bilinguals in Producing the English /a/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	7.65460	0.239	71.967	32.027	< .0001
E mono vs A-E biling	-0.435	0.277	256.057	-1.572	= .117
E mono vs E-A biling	-0.049	0.280	256.110	-0.177	= .860
Speaking condition	0.252	0.273	241.105	0.924	= .356
E mono vs A-E biling x speaking condition	-0.238	0.383	241.145	-0.623	= .534
E mono vs E-A biling x speaking condition	-0.219	0.38517	240.97374	-0.569	= .570

Results for the Z3-Z2 Values of the /a/ Vowel in Arabic and English

Comparison of Monolingual Patterns

In this section, I determined whether A mono and E mono differed in their productions of the /a/ vowel in terms of the Z3-Z2, that is, the degree of frontedness/backness. Figure 3.5 displays the Z3-Z2 values (associated with tongue frontedness/backness) of the Arabic and English /a/ produced by A mono, E mono, E-A bilinguals and A-E bilinguals. To determine whether the Z3-Z2 of the /a/ sound differed across groups, I first compared the Z3-Z2 of both monolingual groups. Figure 3.5 shows that the Z3-Z2 for the E mono speakers was numerically higher than for the A mono speakers, suggesting that E mono tend to produce the /a/ vowel in a less frontal position than do A mono.

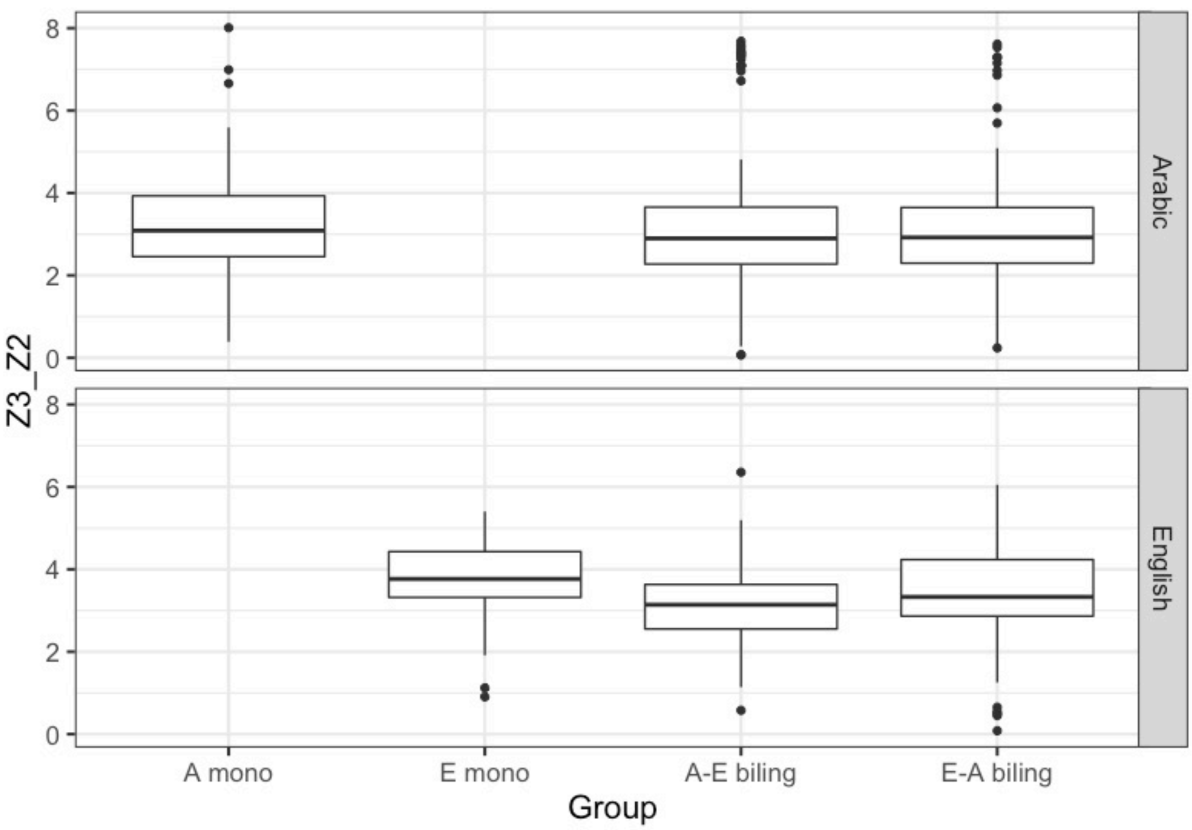


Figure 3.5. Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /a/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

Linear mixed effects models were used to determine whether A mono and E mono differed statistically significantly in terms of the tongue backness/frontedness when producing /a/ (Figure 3.5). Table 3.14 presents the results. The analysis revealed no significant effect of speaker group, showing that the tongue frontedness/backness was similar for E mono and A mono speakers. There was no significant main effect of speaking condition and the speaker group by speaking condition interaction was not significant. Since there was no significant difference in Z3-Z2 values across the two monolingual participant groups, i.e. since the two languages do not differ in terms of tongue frontedness in the production of the /a/ vowel, no analyses for bilingual participants will be conducted.

Table 3.14. Z3-Z2 Comparison between A Mono and E Mono in Producing /a/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	3.157	0.187	65.342	16.808	< .0001
Speaker group	0.334	0.229	170.417	1.458	= .147
Speaking condition	0.154	0.222	156.069	0.695	= .488
Speaker group x speaking condition	0.265	0.3165	155.535	0.840	= .402

3.6.2 Results for the Z3-Z1 Values of the /ɪ/ Vowel in Arabic and English Across Speaker Groups

Next, I consider the short vowel /ɪ/. Figure 3.6 and 3.7 show the mean Z3-Z1 and Z3-Z2 (Bark) frequencies of the short vowel /ɪ/ as produced by Arabic monolinguals, English monolinguals, A-E bilinguals and E-A bilinguals. Z3-Z1 represents the vertical dimension and Z3-Z2 represents in the horizontal dimension. As shown in Figure 3.7, the production of the vowel /ɪ/ is similar in terms of the tongue position across the A-E bilinguals, E-A bilinguals and English monolinguals; Arabic monolinguals and E-A bilinguals differ noticeably from the others. This difference could be caused by outliers in the data, which likely represent one person in each group with a unique production of this particular vowel.

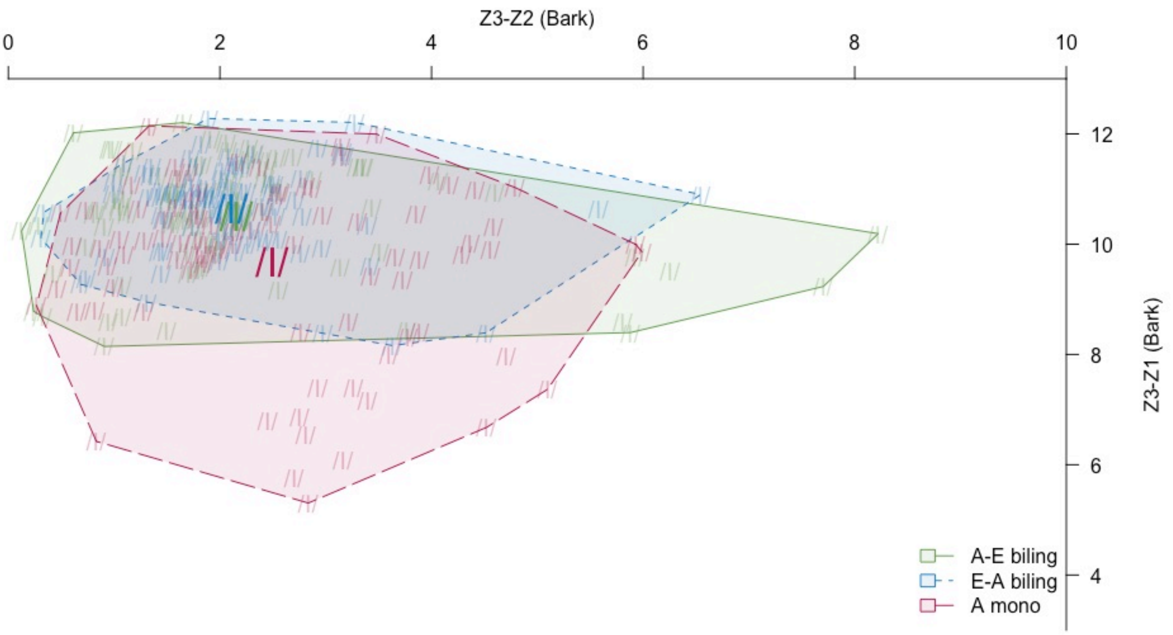


Figure 3.6. Mean Z3-Z1 and Z3-Z2 (Bark) for the Arabic /ɪ/ vowel produced by A mono, A-E biling and E-A biling

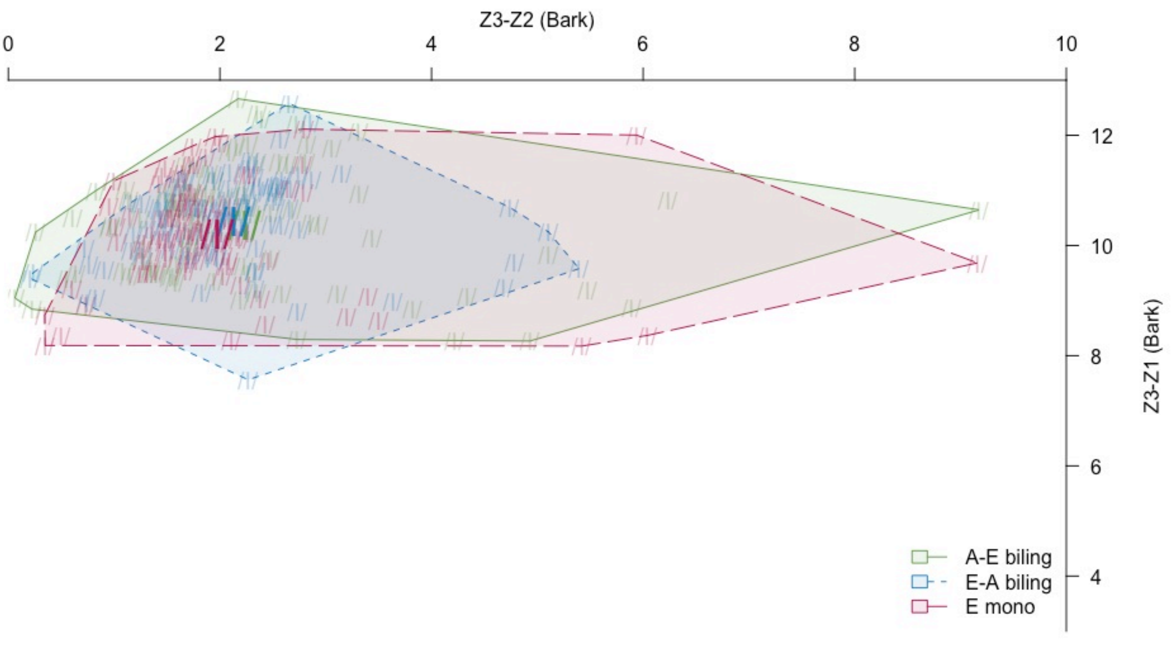


Figure 3.7. Mean Z3-Z1 and Z3-Z2 (Bark) for the English /ɪ/ vowel produced by E mono, A-E biling and E-A biling

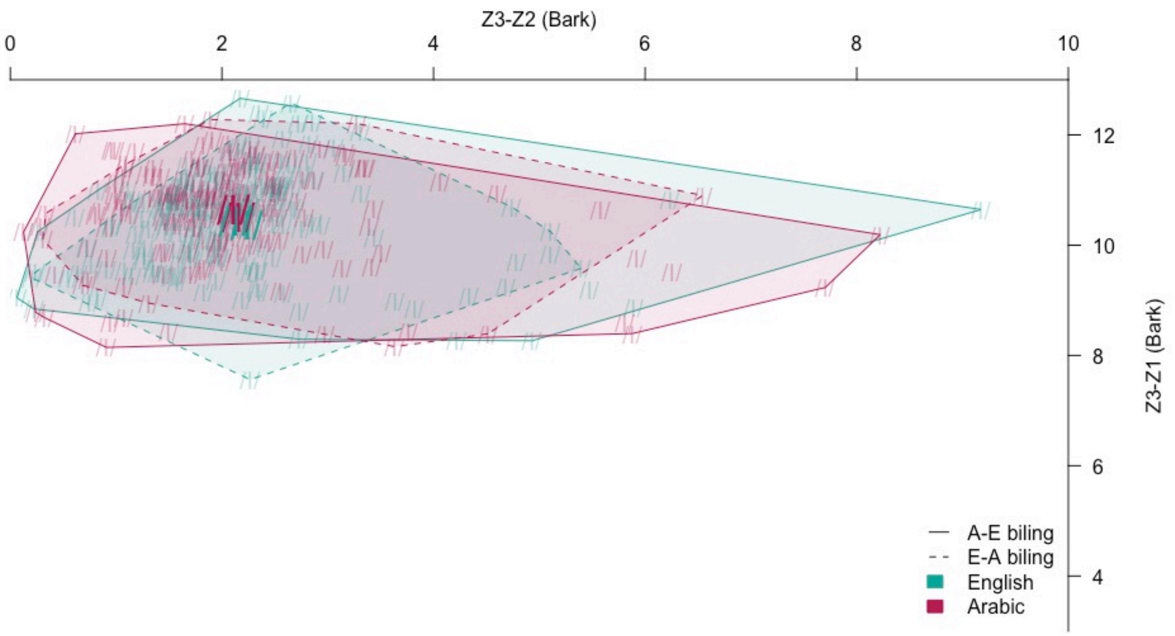


Figure 3.8. Mean Z3-Z1 and Z3-Z2 (Bark) for /ɪ/ vowel produced by A-E biling and E-A biling in Arabic and English

Comparison of Monolingual Patterns

I first compared the E mono and A mono productions of the /ɪ/ vowel in terms of Z3-Z1 to determine whether tongue height differed when producing /ɪ/ in A mono and E mono speakers and to determine whether the productions differed in isolation compared to in a carrier phrase. Figure 3.8 shows the Z3-Z1 values of the Arabic and English /ɪ/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals. As shown in Figure 3.9, the Z3-Z1 is higher numerically for E mono than for A mono speakers, suggesting that E mono are likely to produce the /ɪ/ vowel with a higher tongue position than do A mono speakers.

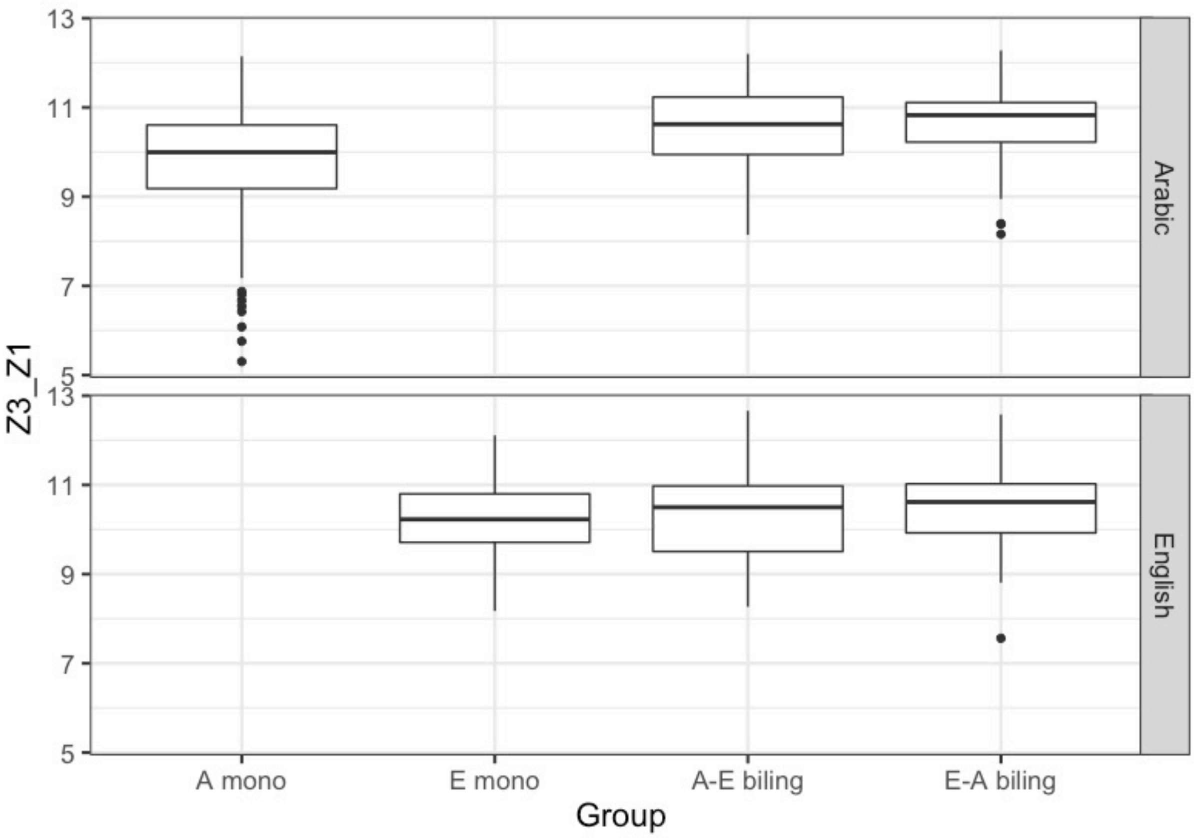


Figure 3.9. Boxplots for the Z3-Z1 values (Bark) of the Arabic and English /ɪ/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

I used linear mixed effects models to determine whether the A mono and E mono speakers shown in Figure 3.9 differ statistically significantly in terms of tongue height of /ɪ/. The significant main effect of the speaker group in Table 3.15 shows a statistically significant difference in tongue height between A mono and E mono speakers in producing /ɪ/, with a significantly higher tongue height for E mono (higher Z3-Z1) than for A mono speakers. In addition, no significant differences were found in the speaking condition. Moreover, the speaker group by speaking condition interaction was not significant.

Table 3.15. Z3-Z1 Comparison between A Mono and E Mono in Producing /ɪ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	9.4719	0.227	40.957	41.673	< .0001
Speaker group	0.8459	0.235	169.730	3.590	< .0001
Speaking condition	0.3213	0.2250	156.832	1.428	= .155
Speaker group x speaking condition	-0.372	0.316	156.533	-1.177	= .241

Bilinguals

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z1 in Producing /ɪ/ in Arabic

Since the Z3-Z1 of the /ɪ/ vowel of the monolingual participants differed, I considered how the productions of bilinguals compared with those of monolingual. First, I compared the bilinguals' Arabic productions with the A mono productions. Figure 3.7 shows the different Z3-Z1 values of the A mono, E-A bilinguals and A-E bilinguals when producing the /ɪ/ vowel in Arabic. I used linear mixed effects models to determine whether the A-E and E-A bilinguals differed statistically significantly from Arabic monolinguals in terms of tongue height when producing /ɪ/. The results revealed that both bilingual groups differed statistically significantly from A mono in terms of tongue height when producing /ɪ/. Thus, neither bilingual group achieved target-like productions in terms of the Z3-Z1 when producing /ɪ/ in Arabic. This finding indicates attrition in the A-E bilinguals' L1. Specifically, the A-E and E-A bilinguals had significantly higher Z3-Z1 values in Arabic, corresponding to a higher tongue position, than the A mono speakers. The A-E and E-A bilinguals' Arabic productions not only moved towards the monolingual English norm, as may be expected in the case of prolonged contact with English for the A-E speakers, but numerically past it, suggesting that the bilinguals may be overshooting the English target in their Arabic productions.

Table 3.16. Z3-Z1 Comparison between A Mono and A-E and E-A Bilinguals in Producing the Arabic /ɪ/.

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	9.523	0.173	126.481	54.784	< .0001
A mono vs A-E biling	0.859	0.2337	248.808	3.677	< .0001
A mono vs E-A biling	1.084	0.233	249.112	4.639	< .0001
Speaking condition	0.318	0.232	247.990	1.371	= .171
A mono vs A-E biling x speaking condition	-0.062	0.328	249.970	-0.189	= .849
A mono vs E-A biling x speaking condition	-0.239	0.329	247.999	-0.726	= .468

Since significant differences exist between the A-E and E-A bilinguals and the monolingual Arabic speakers, I consider whether language use or phonetic aptitude influenced the bilinguals' productions. To determine whether aptitude and language use influenced the Z3-Z1 for /ɪ/ in Arabic produced by the A-E and E-A bilinguals, I used linear mixed effects models. Tables 3.17 and 3.18 summarise the statistical results and show that no significant differences exist between the aptitude groups and language use groups in terms of the Z3-Z1 in the production of /ɪ/ in Arabic.

Table 3.17. Z3-Z1 Comparison between A-E Bilinguals with Low and High Aptitudes and More Arabic and More English in Producing /ɪ/in Arabic.

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	517.85	32.92	11	15.729	< .0001
Aptitude	-29.24	38.02	11	-0.769	= .458
Language use	-46.07	40.32	11	-1.142	= .278
Aptitude x language use	43.74	52.06	11	0.840	= .419

Table 3.18. Z3-Z1 Comparison between E-A Bilinguals with Low and High Aptitudes and More Arabic and More English in Producing /ɪ/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	489.55	32.550	11.87	15.040	< .0001
Aptitude	-4.221	49.141	11.395	-0.09	= .933
Language use	-18.983	45.563	11.454	-0.42	= .694
Aptitude x language use	44.897	66.640	11.198	0.674	= .514

Comparison of A-E and E-A bilinguals with E Monolingual in the Z3-Z1 in producing /ɪ/ in English

Because the Z3-Z1 of the /ɪ/ vowel of the monolingual speakers differed, I also compared the productions of bilinguals with monolingual English speakers. As shown in Figure 3.7, the results showed similar Z3-Z1 values for E mono speakers and the A-E and E-A bilinguals when producing the /ɪ/ vowel in English. I used linear mixed effects models to determine whether E mono and both groups of bilinguals differed statistically significantly when producing /ɪ/ in terms of tongue height. The results showed that neither bilingual group differed from the E mono speakers. As a result, both E-A and A-E bilinguals produced the English /ɪ/ like native monolinguals in terms of the tongue height. Furthermore, speaking condition did not have a significant effect, and the speaker group by speaking condition interaction was not significant.

Table 3.19. Results of the Z3-Z1 Comparison between E Mono and A-E and E-A Bilinguals in Producing the English /ɪ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.203	0.160	55.565	63.566	< .0001
E mono vs A-E biling	0.158	0.185	260.836	0.854	= .394
E mono vs E-A biling	0.273	0.184	260.848	1.485	= .139
Speaking condition	-0.051	0.178	240.882	-0.288	= .773
E mono vs A-E biling x speaking condition	0.055	0.252	240.882	0.221	= .825
E mono vs E-A biling x speaking condition	-0.063	0.251	240.916	-0.254	= .800

Results for the Z3-Z2 Values of the /ɪ/ Vowel in Arabic and English

Comparison of Monolingual Patterns

In this section, I compared the productions of the monolinguals to determine whether A mono and E mono differed in producing the /ɪ/ vowel in terms of Z3-Z2 , that is, in terms of the degree of backness/frontedness, and to determine whether their productions of the vowel differed in isolation compared to in a carrier phrase. Figure 3.10 shows the Z3-Z2 values (associated with tongue frontedness/backness) of the Arabic and English /ɪ/ produced by A mono, E mono and the A-E and E-A bilinguals. Figure 3.10 shows that the Z3-Z2 of E mono was numerically lower than that of A mono, indicating that E mono tend to produce the /ɪ/ vowel with a more frontal tongue position than do A mono speakers.

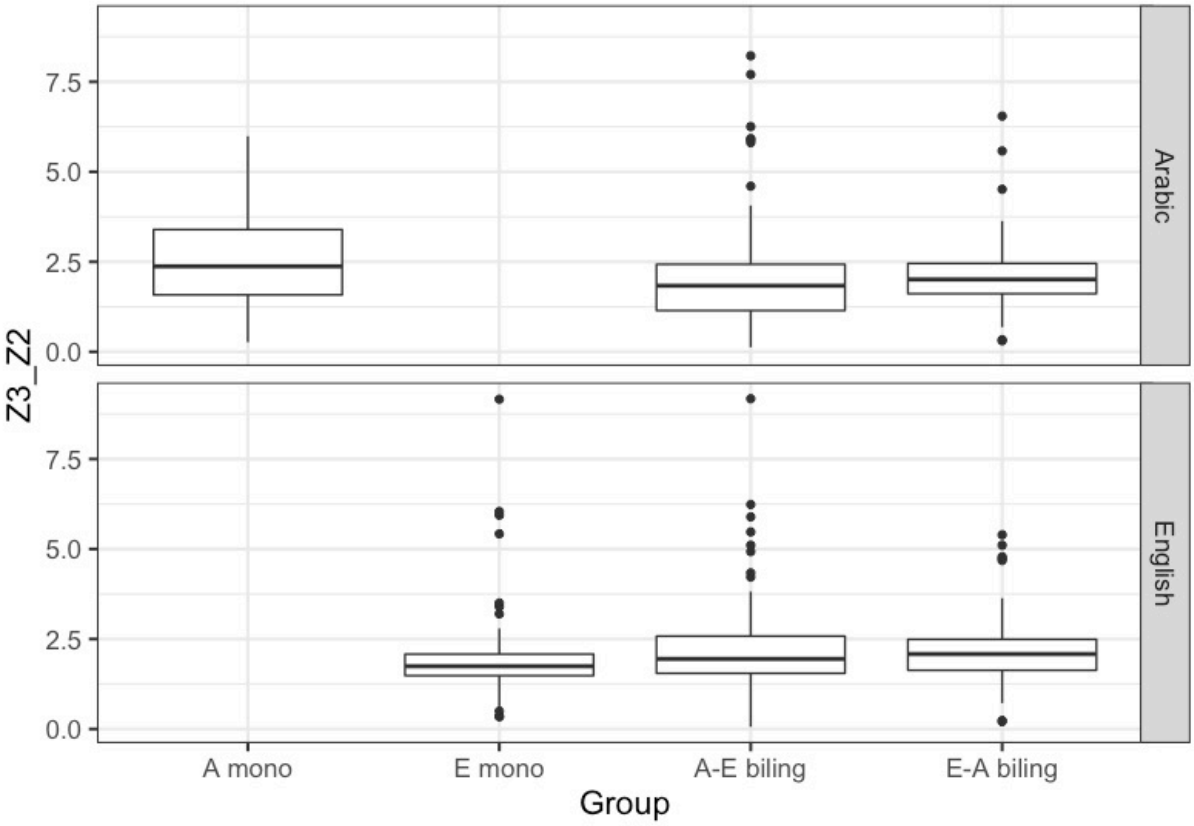


Figure 3.10. Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /ɪ/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

Linear mixed effects models were used to determine whether the monolingual control groups differed statistically significantly in terms of tongue backness/frontedness when producing /ɪ/ (Figure 3.10). Table 3.20 presents the results. The analysis showed that speaker group had a significant effect, showing that the tongue was significantly more fronted for E mono than for A mono. There was no main effect of speaker group and the speaker group by speaking condition interaction was not significant.

Table 3.20. Results of the Z3-Z2 Comparison between A Mono and E Mono in Producing the Arabic /ɪ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	2.662	0.253	30.585	10.484	< .0001
Speaker group	-0.982	0.233	166.171	-4.206	< .0001
Speaking condition	0.155	0.221	154.844	0.703	= .483
Speaker group x speaking	0.1047	0.311	154.610	0.336	= .737

condition

Bilinguals

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z2 in Producing /ɪ/ in Arabic

I examined how bilinguals' productions compared with those of monolinguals because the monolingual participants differed in producing the /ɪ/ vowel in terms tongue frontedness. I first compared A mono productions with bilinguals' Arabic productions. As shown in Figure 3.10, A mono, A-E bilinguals and E-A bilinguals numerically differed in the production of /ɪ/ in Arabic. I used linear mixed effects models to determine whether the E-A and A-E bilinguals displayed statistically significant differences in terms of tongue backness/frontedness when producing /ɪ/. The results revealed that neither bilingual group differed significantly from A mono. Thus, both of the bilingual groups achieved native-like productions in terms of frontedness for the Arabic vowel /ɪ/.

Table 3.21. Z3-Z2 Comparison between A Mono and A-E and E-A Bilinguals in Producing the Arabic /ɪ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	2.416	0.216	59.620	11.150	< .0001
A mono vs A-E biling	-0.341	0.253	247.370	-1.348	= .179
A mono vs E-A biling	-0.255	0.253	247.640	-1.007	= .315

Comparison of A-E and E-A Bilinguals with E Monolingual in the Z3-Z2 in Producing /ɪ/ in English

As the Z3-Z2 of the /ɪ/ vowel differed in terms of A mono and E mono productions, I compared the bilinguals' English productions with those of the English monolinguals. Figure 3.10 shows numerically different Z3-Z2 values for the E mono and E-A bilingual productions of /ɪ/, whereas the A-E bilinguals were similar to E mono in their L2 productions in terms of frontedness. I used linear mixed effects models to determine whether E mono speakers and the E-A and A-E bilinguals differed statistically significantly in terms of tongue backness/frontedness when producing the English /ɪ/. The results revealed that the A-E bilinguals did not differ from E mono in this regard. However, the E-A bilinguals differed statistically significantly from E mono. The Z3-Z2 for E-A bilinguals is higher than for E mono, suggesting that the E-A bilinguals produced the English /ɪ/ with a more fronted tongue position than E mono. As a result, the A-E bilinguals did achieve native-like productions of the English /ɪ/ in terms of frontedness. In contrast, the E-A bilinguals showed evidence of attrition, with values in the direction of the Arabic norm.

Table 3.22. Z3-Z2 Comparison between E Mono and A-E and E-A Bilinguals in Producing the English /ɪ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	1.806	0.217	41.630	8.319	< .0001
E mono vs A-E biling	0.224	0.229	258.210	0.975	= .330
E mono vs E-A biling	0.587	0.228	258.220	2.567	= .010

Since a significant difference exists between the E-A bilinguals and the monolingual English speakers, I considered the language use and phonetic aptitude factors. I used linear mixed effects models to determine whether aptitude and language use influenced the E-A bilinguals' productions of English /ɪ/ in terms of frontedness (Z3-Z2). The statistical results summarised in Table 3.23 show that no significant differences exist between the aptitude and language use groups in terms of the Z3-Z2 for the English /ɪ/.

Table 3.23. Z3-Z2 Comparison between E-A Bilinguals with Low and High Aptitude and More Arabic and More English in Producing /ɪ/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	529.435	32.645	11.267	16.218	< .0001
Aptitude	1.067	43.802	13.550	0.024	= .981
Language use	-51.417	46.167	11.267	1.114	= .289
Aptitude x language use	47.591	54.970	17.657	0.866	= .398

3.6.3 Results for the Z3-Z1 Values of the /ʊ/ Vowel in Arabic and English Across Speaker Groups

In this section, I consider the short vowel /ʊ/. Figure 3.11 and 3.12 shows the mean Z3-Z1 and Z3-Z2 values (in Bark) of the short vowel /ʊ/ that were produced by Arabic monolinguals, English monolinguals and the A-E and E-A bilinguals. Z3-Z1 represents the vertical dimension and Z3-Z2 represents in the horizontal dimension. As with the /ɪ/ vowel productions, a close similarity seems to exist between all groups of participants in terms of their tongue positions. However, every group has a number of outliers, but it is not immediately clear why this is the case.

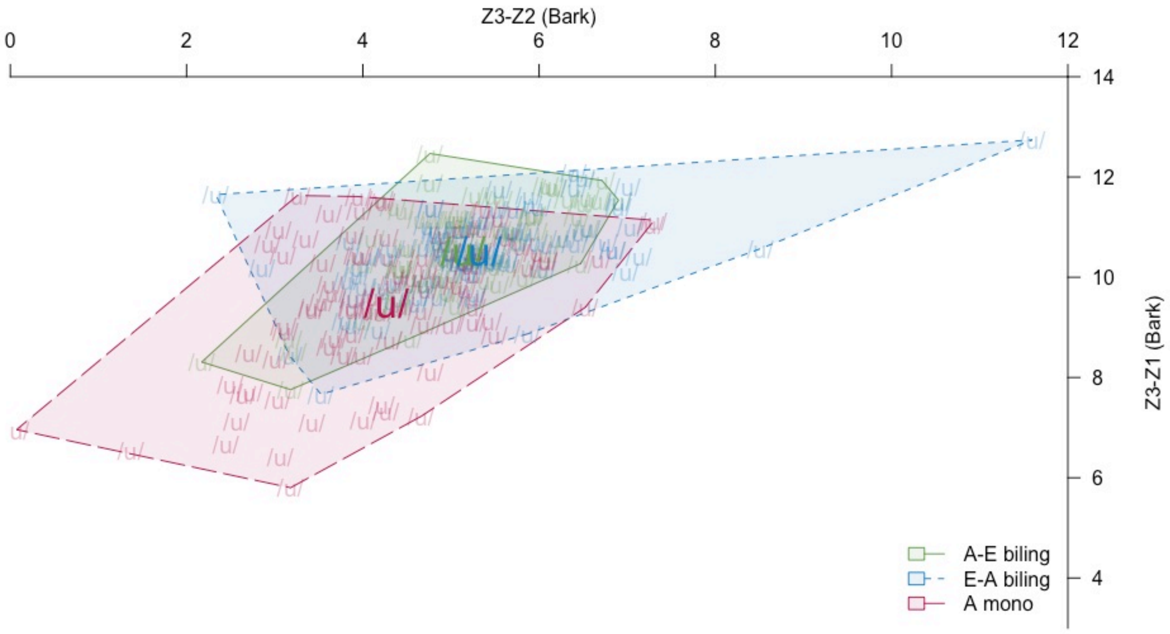


Figure 3.11. Mean Z3-Z1 and Z3-Z2 for the Arabic /ʊ/ vowel produced by A mono, A-E biling and E-A biling

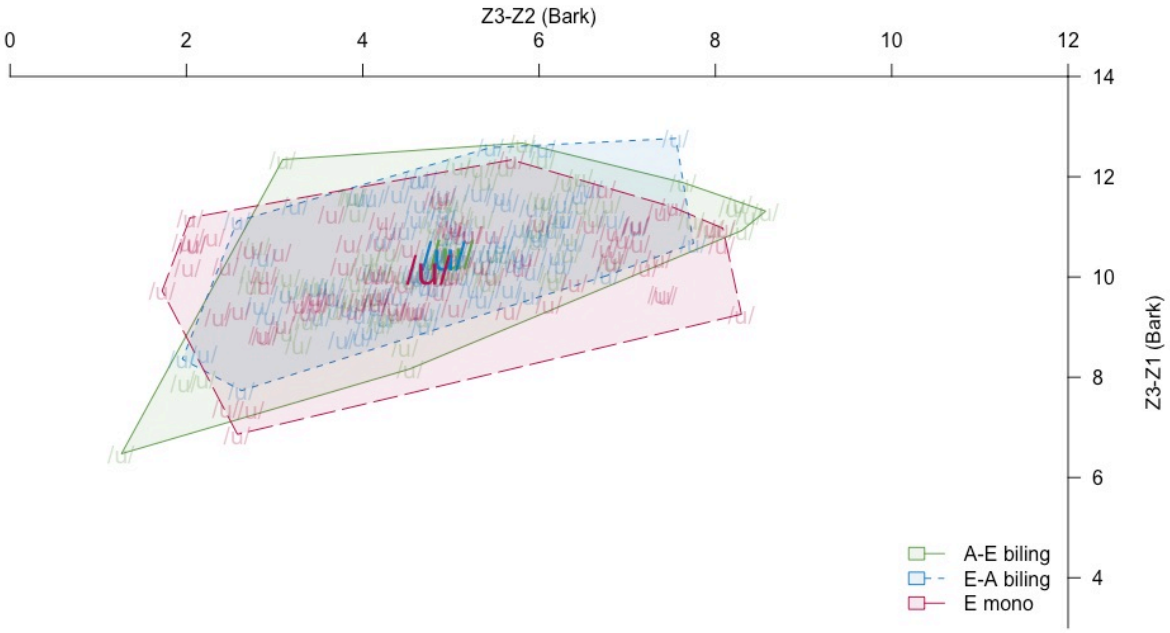


Figure 3.12. Mean Z3-Z1 and Z3-Z2 for the English /ʊ/ vowel produced by E mono, A-E biling and E-A biling

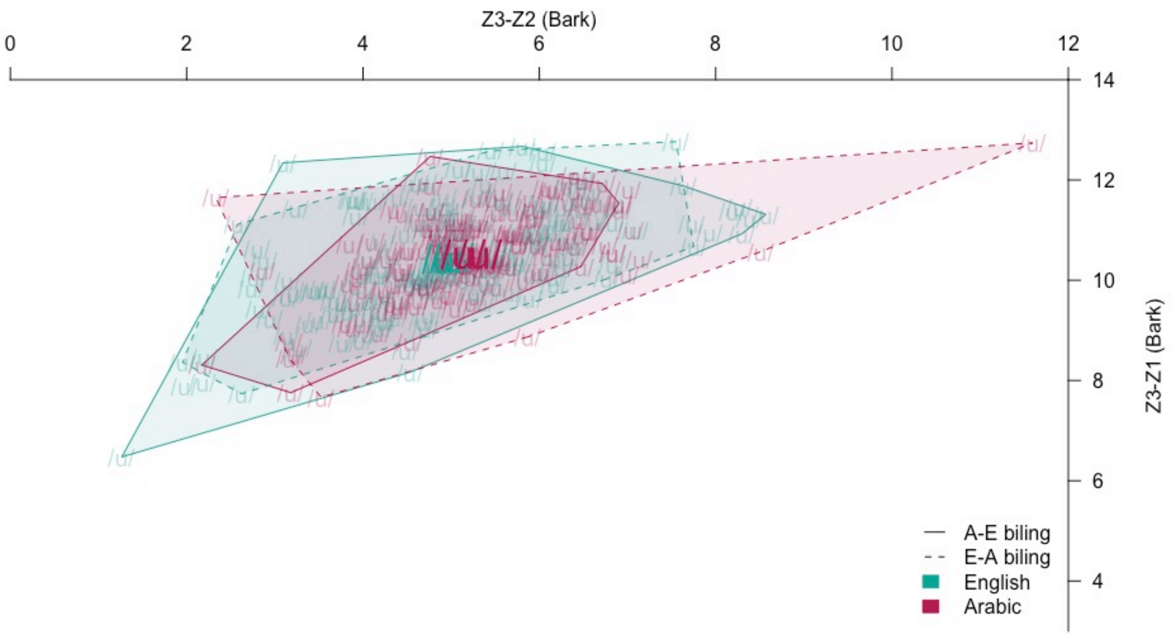


Figure 3.13. Mean Z3-Z1 and Z3-Z2 for /ʊ/ vowel produced by Amono, E mono, A-E biling and E-A biling

Comparison of Monolingual Patterns

First, I compared the productions of the /ʊ/ vowel in terms of Z3-Z1 (tongue height) in E mono and A mono speakers. Figure 3.14 shows the Z3-Z1 values of the Arabic and English /ʊ/ as produced by the A-E bilinguals, E-A bilinguals and the A mono and E mono speakers. To determine whether tongue height differed between the two languages when producing the /ʊ/ vowel, I compared the Z3-Z1 of the A mono and E mono speakers. Figure 3.14 shows that the Z3-Z1 of E mono was a slightly higher than A mono, suggesting that E mono speakers tend to produce the /ʊ/ vowel with a higher tongue height than the A mono.

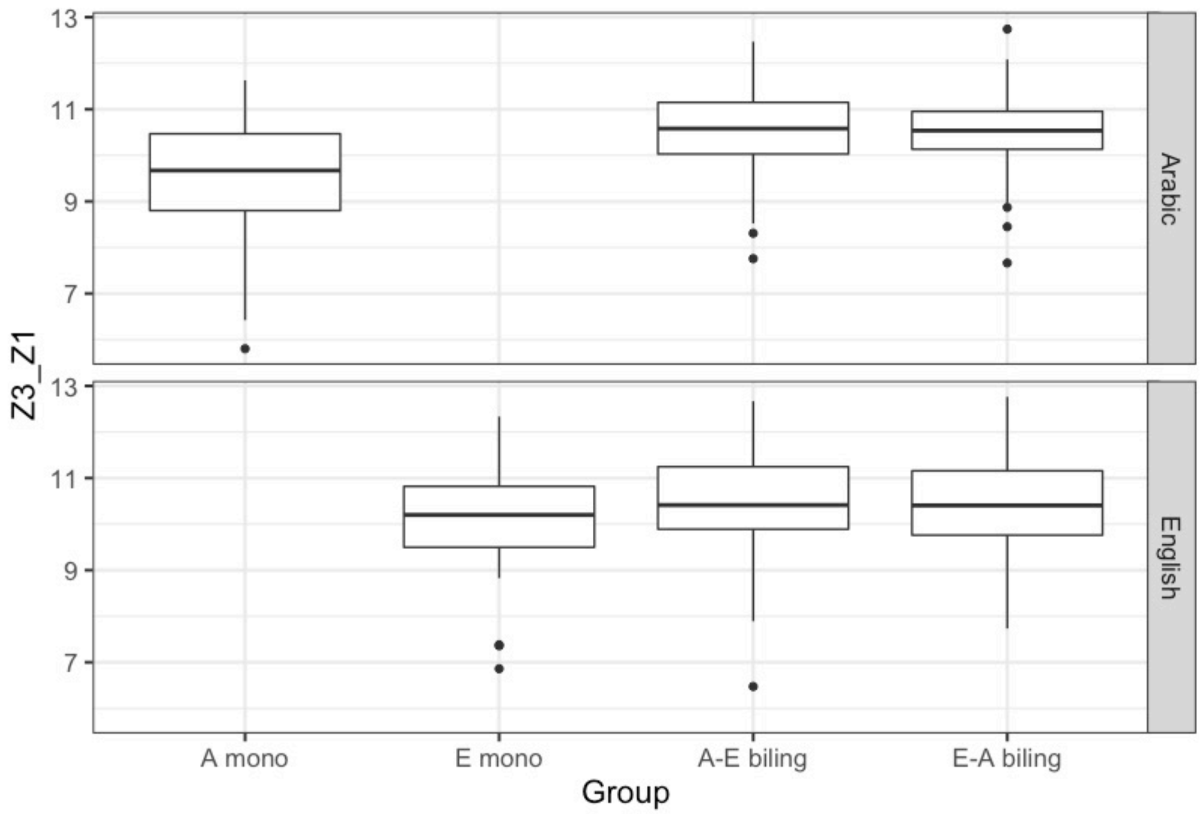


Figure 3.14. Boxplots for the Z3-Z1 values (Bark) of the Arabic and English /ʊ/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

I used linear mixed effects models to determine whether Z3-Z1 differed statistically significantly in the productions of /ʊ/ across the monolingual participant groups and in isolation compared to in a carrier phrase. Table 3.24 presents the results. The analysis shows that E mono produced /ʊ/ with significantly higher tongue position than A mono. Moreover, speaking condition just failed to reach significance, and the speaker group by speaking condition interaction was not significant.

Table 3.24. Results of the Z3-Z1 Comparison between A Mono and E Mono in Producing /ʊ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	9.300	0.199	49.720	46.591	< .0001
Speaker group	0.837	0.225	174.140	3.718	< .0001
Speaking condition	0.416	0.217	159.530	1.913	= .057
Speaker group x speaking condition	-0.237	0.307	159.400	-0.772	= .441

Bilinguals

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z1 in Producing /ʊ/ in Arabic

Since the Z3-Z1 of the /ʊ/ vowel of the monolingual participants differed, I considered how the productions of bilinguals compared with those of monolinguals. First, I compared the productions of the bilinguals in Arabic with the A mono productions. Figure 3.14 displays the different Z3-Z1 values of the A mono, E-A bilinguals and A-E bilinguals when producing the Arabic /ʊ/ vowel. I used linear mixed effects models to determine whether the A-E and E-A bilinguals differed statistically significantly from Arabic monolinguals in terms of tongue height when producing /ʊ/. The results revealed that both bilingual groups differed statistically significantly from A mono in terms of tongue height when producing /ʊ/. Thus, neither bilingual group achieved target-like productions of the Arabic /ʊ/ in terms of tongue height. This finding indicates attrition in the A-E bilinguals' L1. In particular, the E-A and A-E bilinguals had significantly higher Z3-Z1 values in Arabic, corresponding to a higher tongue position, than the A mono.

Table 3.25. Z3-Z1 Comparison between A Mono and A-E and E-A Bilinguals in Producing the Arabic /ʊ/.

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	9.290	0.172	63.880	53.722	< .0001
A mono vs A-E biling	0.205	1.220	252.920	5.939	< .0001
A mono vs E-A biling	1.164	0.201	253.450	5.792	< .0001
Speaking condition	0.434	0.202	252.440	2.142	= .194
A mono vs A-E biling x speaking condition	-0.386	0.288	253.510	-1.340	= .181
A mono vs E-A biling x speaking condition	-0.233	0.287	252.870	-0.814	= .416

Because significant differences exist between the A-E and E-A bilinguals and the monolingual Arabic speakers, I consider whether language use or phonetic aptitude influenced the bilinguals' productions. I used linear mixed effects models, to determine whether aptitude and language use influenced the Z3-Z1 for /ʊ/ in Arabic produced by the A-E and E-A bilinguals. Tables 3.26 and 3.27 summarise the statistical results and show that no significant differences exist between the aptitude groups and language use groups in terms of the Z3-Z1 in the production of /ʊ/ in Arabic.

Table 3.26. Z3-Z1 Comparison between A-E Bilinguals with Low and High Aptitudes and More Arabic and More English in Producing /ʊ/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.521	0.547	11	19.221	< .0001
Aptitude	-0.269	0.632	11	-0.427	= .678
Language use	0.587	0.670	11	0.876	= .400
Aptitude x language use	-0.485	0.865	11	-0.562	= .586

Table 3.27. Z3-Z1 Comparison between E-A Bilinguals with Low and High Aptitudes and More Arabic and More English in Producing /ʊ/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.483	0.303	11.072	34.558	< .0001
Aptitude	0.305	0.460	10.832	0.663	= .521
Language use	0.317	0.426	10.862	0.743	= .473
Aptitude x language use	-0.849	0.625	10.670	-1.359	= .202

Comparison of A-E and E-A bilinguals with E Monolingual in the Z3-Z1 in producing /ʊ/ in English

I also compared the English productions of bilinguals with monolingual English speakers, since the Z3-Z1 of the /ʊ/ vowel of the monolingual speakers differed. As shown in Figure 3.14, the results showed similar Z3-Z1 values for E mono speakers and the A-E and E-A bilinguals when producing the /ʊ/ vowel in English. I used linear mixed effects models to determine whether E mono differed statistically significantly from A-E and E-A bilinguals when producing /ʊ/ in terms of tongue height. The A-E bilinguals did not differ from the E mono speakers, suggesting target-like productions in terms of tongue height. In contrast, E-A bilinguals produced /ʊ/ with significantly higher tongue than E monolinguals. This finding indicates attrition in the E-A bilinguals' L1. Furthermore, speaking condition did not have a significant effect, and the speaker group by speaking condition interaction was not significant.

Table 3.28. Results of the Z3-Z1 Comparison between E Mono and A-E and E-A Bilinguals in Producing the English /ʊ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.066	0.163	82.080	61.594	< .0001
E mono vs A-E biling	0.139	0.205	261.460	0.678	= .498
E mono vs E-A biling	0.427	0.204	260.960	2.093	= .037
Speaking condition	0.164	0.198	241.000	0.830	= .407
E mono vs A-E biling x speaking condition	0.319	0.283	241.850	1.128	= .260
E mono vs E-A biling x speaking condition	-0.2753	0.281	240.950	-0.979	= .328

Since significant differences exist between the E-A bilinguals and the monolingual English speakers, I consider whether language use or phonetic aptitude influenced the bilinguals' productions. To determine whether aptitude and language use influenced the Z3-Z1 for /ʊ/ in English produced by the E-A bilinguals, I used linear mixed effects models. Table 3.29 summarises the statistical results and show that no significant differences exist between the aptitude groups and language use groups in terms of the Z3-Z1 in the production of /ʊ/ in English.

Table 3.29. Z3-Z1 Comparison between E-A Bilinguals with Low and High Aptitudes and More Arabic and More English in Producing /ʊ/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.385	0.366	11	28.315	< .0001
Aptitude	0.579	0.560	11	1.035	= .323
Language use	-0.053	0.518	11	0.103	= .-920
Aptitude x language use	-0.730	0.763	11	-0.956	= .359

Results for the Z3-Z2 Values of the /ʊ/ Vowel in Arabic and English

Comparison of Monolingual Patterns

Figure 3.15 shows the Z3-Z2 of the Arabic and English /ʊ/ produced by the E-A and A-E bilinguals and the A mono and E mono speakers. First, the Z3-Z2 values of both monolingual groups were compared. The figure shows that the Z3-Z2 values for E mono were numerically higher than for A mono. The results of the linear mixed effects model, reported in Table 3.30, shows that the Z3-Z2 values for E mono compared to A mono just failed to reach significance. There was also no significant effect of speaking condition and the speaker group by speaking condition interaction.

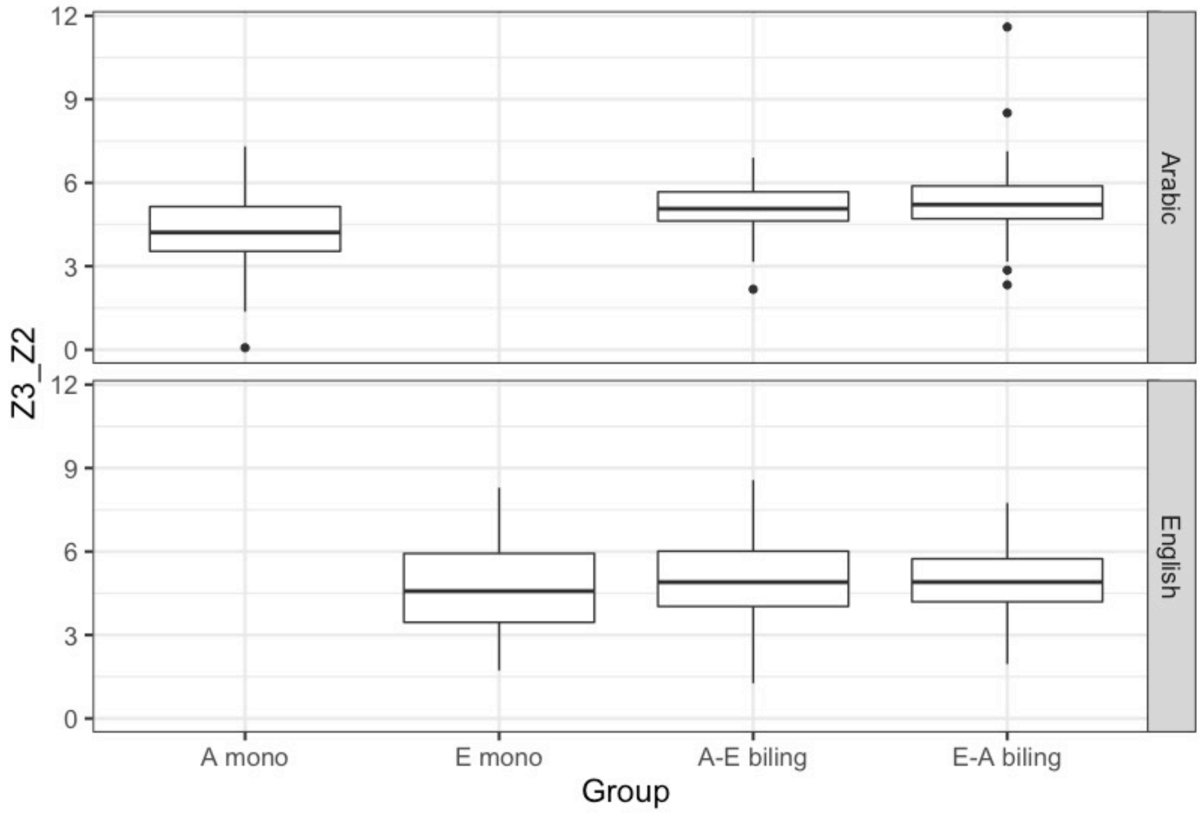


Figure 3.15 Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /ʊ/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

Table 3.30. Results of the Z3-Z2 Comparison between A Mono and E Mono in Producing /ʊ/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	4.2073	0.296	29.600	14.172	< .0001
Speaker group	0.473	0.245	167.770	1.929	= .055

Speaking condition	0.269	0.232	159.650	1.158	= .248
Speaker group x speaking condition	0.249	0.328	159.610	0.759	= .448

Since monolingual participants did not differ in the Z3-Z2 of their /ʊ/ productions, I did not compare the bilinguals' productions of the /ʊ/ vowel in terms of Z3-Z2 with those of the monolinguals.

3.6.4 Long Vowels

Results for the Z3-Z1 Values of the /ɑː/ Vowel in Arabic and English Across Speaker Groups

After considering the short vowels, I move on to the long vowels, beginning with the long vowel /ɑː/. Figure 3.16 and 3.17 show the mean Z3-Z1 and Z3-Z2 values (in Bark) of the long vowel /ɑː/ produced by Arabic monolinguals, English monolinguals, A-E bilinguals and E-A bilinguals. Z3-Z1 represents the vertical dimension and F2 represents the horizontal dimension. Regarding tongue height, no difference seems to exist across any of the groups (except for a small number of outliers in the productions of the Arabic monolinguals). However, when tongue frontedness is considered, the English productions of the /ɑː/ vowel are less fronted than are those of the Arabic productions.

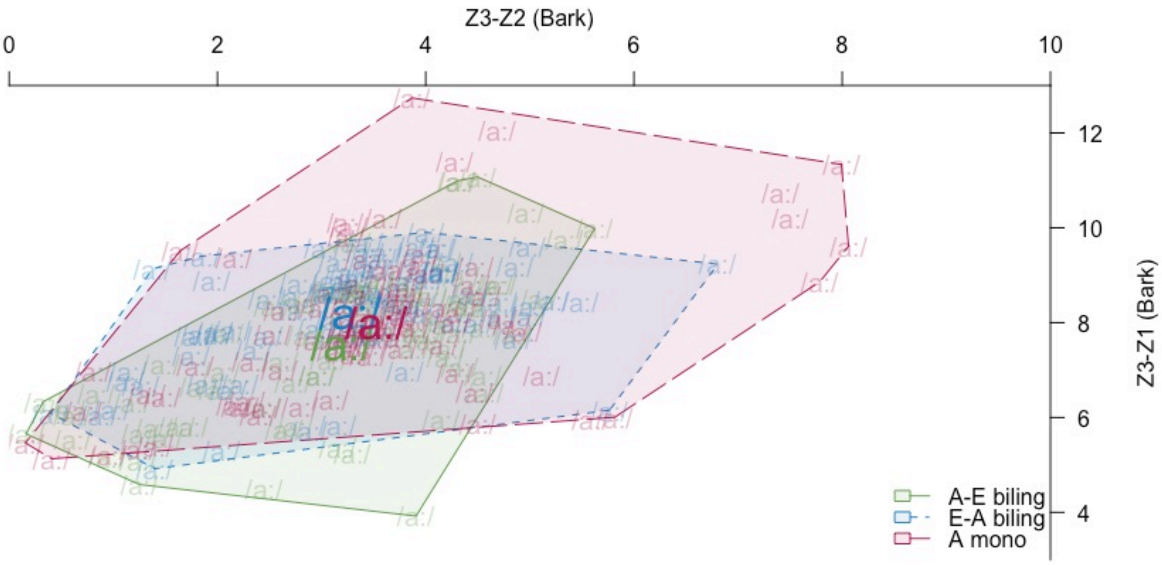


Figure 3.16. Mean Z3-Z1 and Z3-Z2 for in Arabic /a:/ vowel produced by A mono, A-E biling and E-A biling

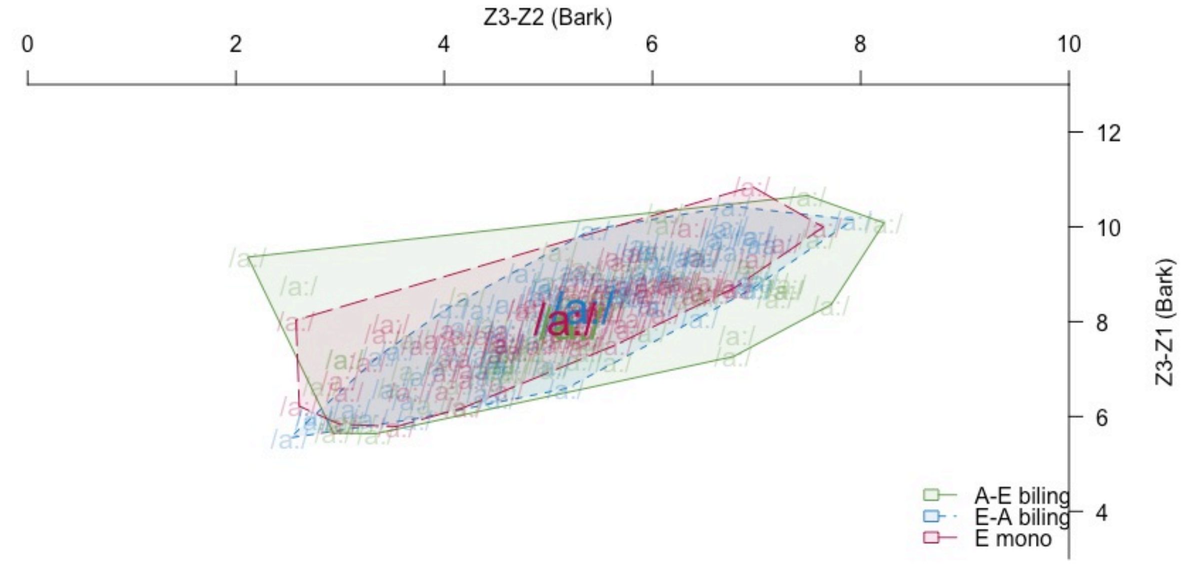


Figure 3.17. Mean Z3-Z1 and Z3-Z2 for the English /a:/ vowel produced by E mono, A-E biling and E-A biling

.

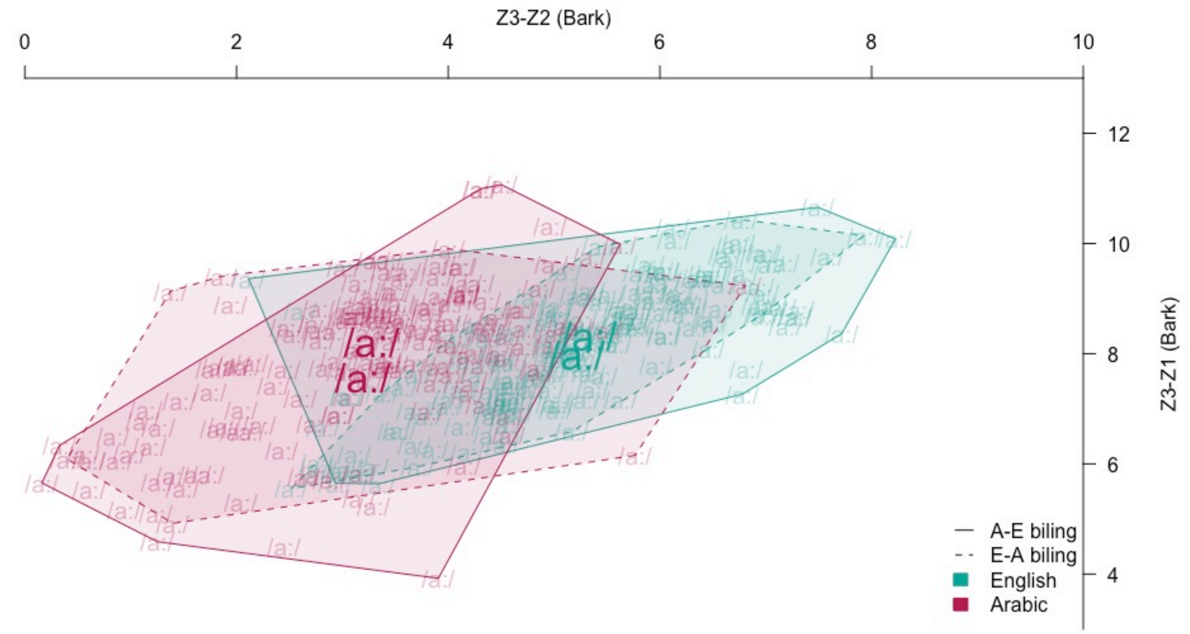


Figure 3.18. Mean Z3-Z1 and Z3-Z2 for /a:/ vowel produced by A-E biling and E-A biling in Arabic and English

.

Comparison of Monolingual Patterns

I determined first whether the A mono and E mono speakers differed statistically significantly in their productions of /ɑː/ in terms of tongue height, i.e. Z3-Z1. Figure 3.19 displays the Z3-Z1 values of the Arabic and English /ɑː/ for A mono, E mono and the A-E and E-A bilinguals. First, I compared the Z3-Z1 for A mono and E mono speakers to determine whether tongue height differed between the two groups in producing the /ɑː/ sound. Figure 3.19 shows that the Z3-Z1 for E mono was numerically similar to that of A mono, suggesting that E mono and A mono speakers tend to produce the /ɑː/ vowel with a similar tongue height.

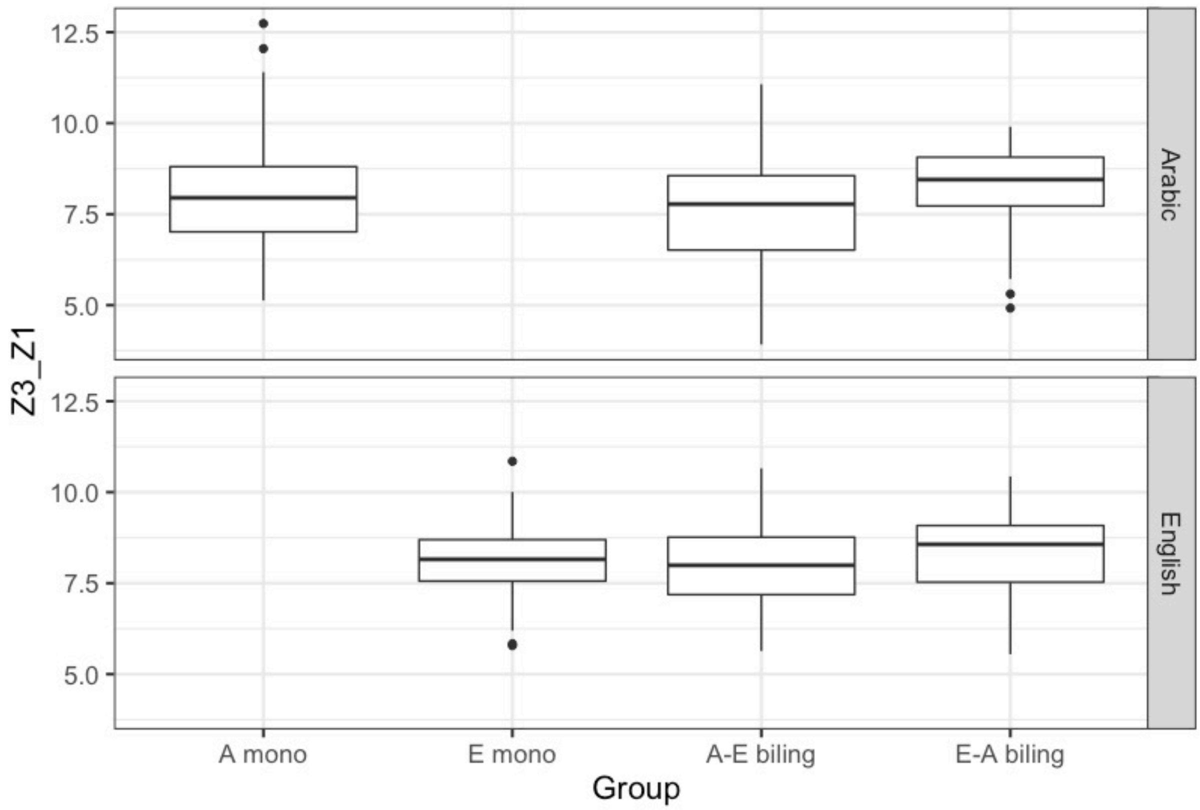


Figure 3.19. Boxplots for the Z3-Z1 values (Bark) of the Arabic and English /ɑː/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

I used linear mixed effects models to determine whether A mono and E mono differed statistically significantly in terms of tongue height and in isolation compared to in a carrier phrase when producing the /ɑː/ sound. Table 3.31 shows the results. No significant difference was evident between E mono and A mono in terms of tongue height. Moreover, speaking condition had no significant effect on the production of the /ɑː/ sound. Finally, the speaker group by speaking condition interaction was not significant.

Table 3.31. Results of the Z3-Z1 Comparison between A Mono and E Mono in Producing /ɑː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	7.844	0.221	54.140	35.381	< .0001
Speaker group	0.204	0.249	173.840	0.820	= .414
Speaking condition	0.259	0.238	159.930	1.088	= .278
Speaker group x speaking condition	0.4796	0.3329	159.640	1.441	= .152

Since monolingual participants did not differ in the Z3-Z1 of their /ɑː/ productions, I did not compare the bilinguals' productions of the /ɑː/ vowel in terms of Z3-Z1 with those of the monolinguals.

Results for the Z3-Z2 Values of the /ɑː/ Vowel in Arabic and English

Comparison of Monolingual Patterns

Figure 3.20 displays the Z3-Z2 of the Arabic and English /ɑː/ vowel produced by A mono, E mono, and the A-E and E-A bilinguals. The Z3-Z2 of E mono was numerically higher than that of A mono, indicating that E mono speakers tend to produce the /ɑː/ vowel with a less frontal tongue position than do A mono speakers.

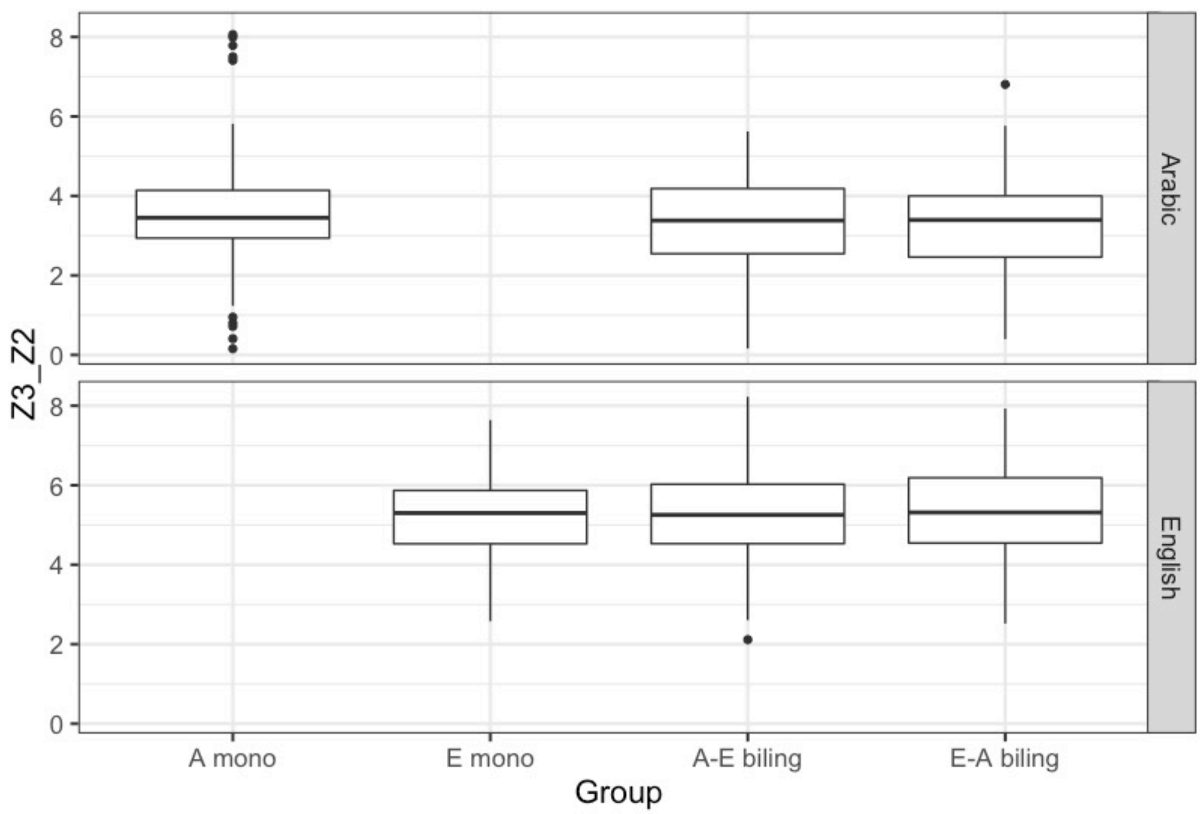


Figure 3.20. Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /ɑː/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

To determine whether A mono and E mono speakers differed statistically significantly in terms of the tongue backness/frontedness when producing /ɑː/, I used linear mixed effects models. Table 3.32 presents the results. The analysis revealed that E mono speakers produced /ɑː/ with a significantly less fronted tongue than did A mono speakers. There was no main effect of speaking condition and the speaker group by speaking condition interaction was not significant.

Table 3.32. Z3-Z2 Comparison between A Mono and E Mono in Producing /ɑː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	5.248	0.234	95.140	22.413	< .0001
Speaker group	-0.972	0.253	266.930	-3.841	< .0001
Speaking condition	0.004	0.275	250.010	0.018	= .985
Speaker group x speaking condition	0.033	0.339	250.210	0.098	= .921

Bilinguals

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z2 in Producing /ɑː/ in Arabic

Since the monolinguals' production of the /ɑː/ vowel differed in terms of the Z3-Z2, I compared their productions of /ɑː/ with those of the bilinguals in terms of Z3-Z2. I first compared the A mono productions with both bilinguals' Arabic productions. Figure 3.20 shows similar Z3-Z2 values across A mono and the bilingual participants. To determine whether the E-A and A-E bilinguals displayed any statistically significant differences when compared to A mono speakers, I used linear mixed effects models. Table 3.33 shows that neither the E-A bilinguals nor the A-E bilinguals differed significantly in terms of Z3-Z2 from A mono. Thus, both the A-E bilingual and the E-A bilingual group had native-like productions in terms of tongue height in Arabic.

Table 3.33. Z3-Z2 Comparison between A Mono and A-E and E-A Bilinguals in Producing /ɑː/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	3.343	0.282	114.270	11.836	< .0001
A mono vs A-E biling	-0.218	0.358	249.050	-0.610	= .542
A mono vs E-A biling	0.270	0.365	249.950	0.742	= .459

Comparison of E-A and A-E Bilinguals with E Monolingual in the Z3-Z2 in Producing /ɑː/ in English

As the monolingual participants' productions of the /ɑː/ vowel differed in the Z3-Z2, I compared the bilinguals' English productions with those of the English monolinguals. Figure 3.20 shows that E mono, A-E bilinguals and E-A bilinguals were similar in the production of /ɑː/ in English in terms of the Z3-Z2. To determine whether the E-A and A-E bilinguals displayed any statistically significant differences from E mono speakers, I used linear mixed effects models. Table 3.34 shows that both the E-A and A-E bilinguals did not differ statistically significantly from the E mono group in terms of Z3-Z2 values. Thus, both groups of bilinguals had native-like productions in terms of tongue frontedness in English.

Table 3.34. Z3-Z2 Comparison between E Mono and A-E and E-A Bilinguals in Producing /a:/ in English

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	8.203	0.175	91.690	46.875	< .0001
E mono vs A-E biling	-0.239	0.213	266.220	-1.119	= .264
E mono vs E-A biling	0.091	0.215	266.360	0.427	= .670

3.6.5 Results for the Z3-Z1 Values of the /iː/ Vowel in Arabic and English Across Speaker Groups

Next, I consider the long vowel /iː/. Figure 3.21 shows the mean Z3-Z1 and Z3-Z2 values (in Bark) of the long vowel /iː/ produced by the Arabic monolinguals, English monolinguals, A-E bilinguals and E-A bilinguals. Z3-Z1 represents the vertical dimension and Z3-Z2 represents the horizontal dimension. When considering tongue height, all groups are similar, with little spread in terms of tongue height, except for the Arabic monolingual speakers, which have a greater range of tongue heights (Z3-Z1) compared to the other groups. However, there is a considerable spread of values in terms of tongue frontedness/backness.

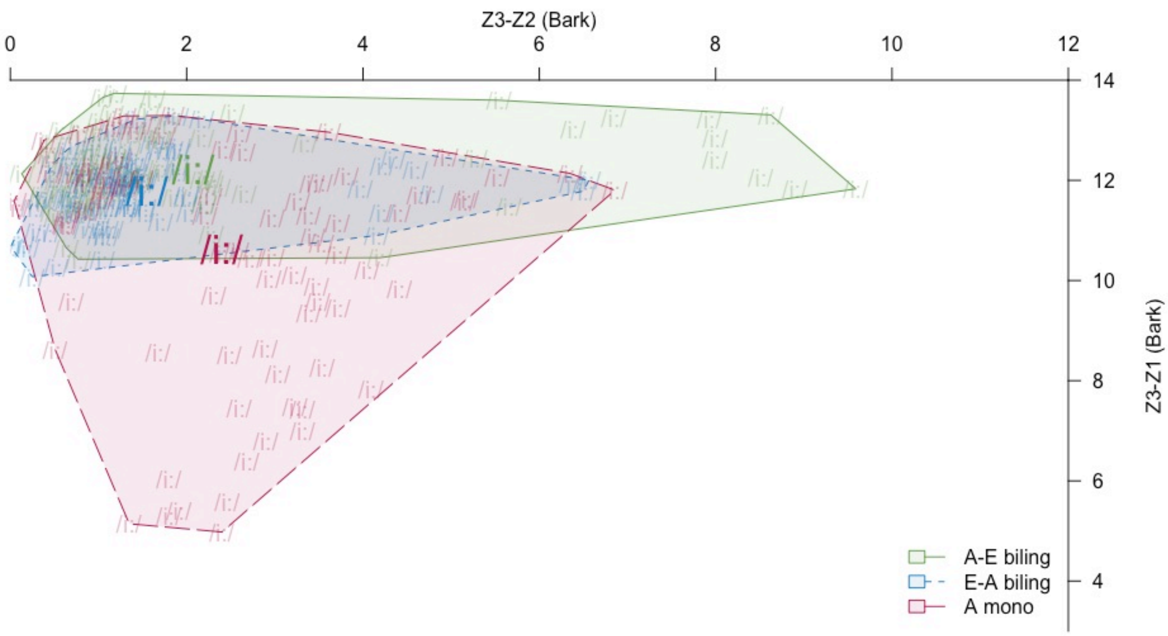


Figure 3.21. Mean Z3-Z1and Z3-Z2 for the Arabic /i:/vowel produced by Amono, A-E biling and E-A biling

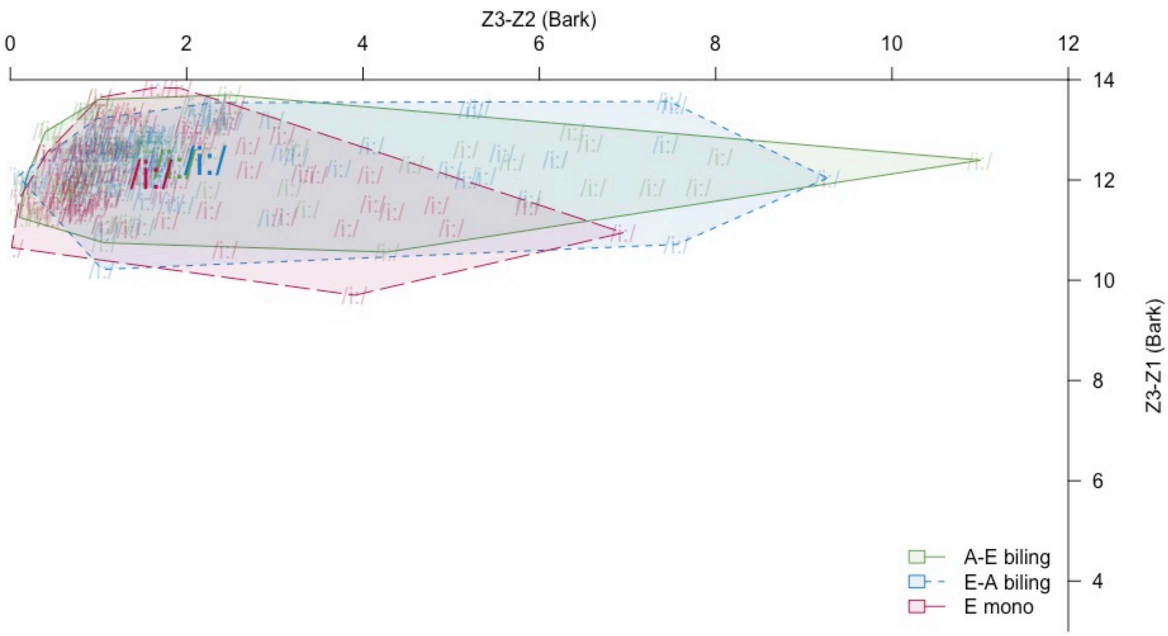


Figure 3.22. Mean Z3-Z1and Z3-Z2 for the English /i:/vowel produced by Emono, A-E biling and E-A biling

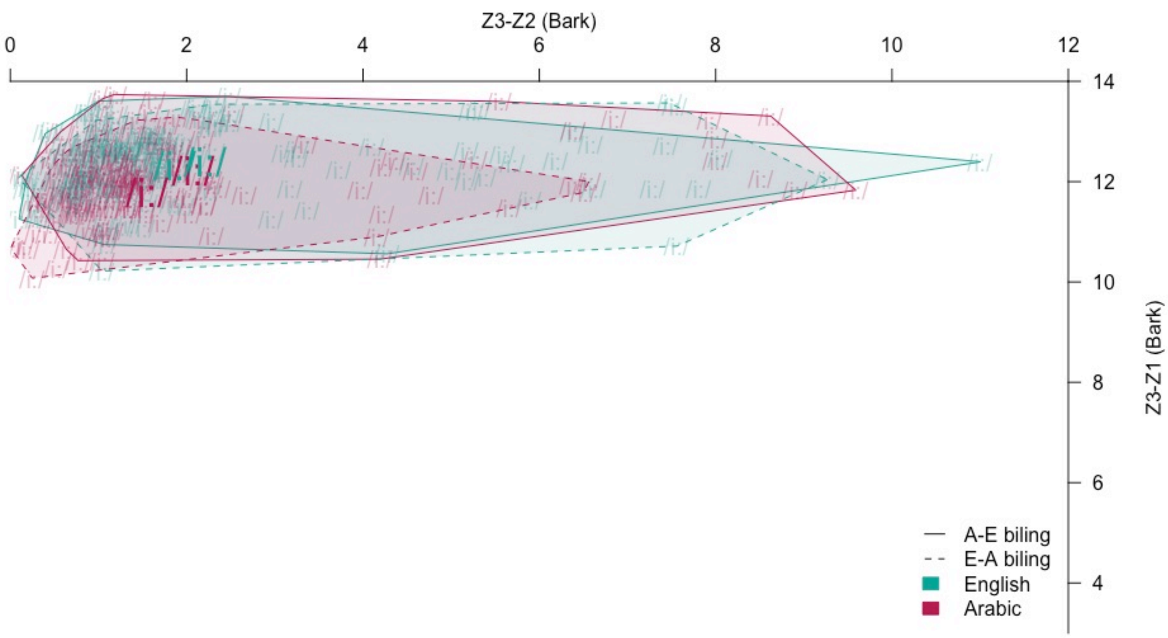


Figure 3.23. Mean Z3-Z1and Z3-Z2 for /i:/vowel produced by A-E biling and E-A biling in both languages

Comparison of Monolingual Patterns

A mono and E mono productions of the /i:/ vowel were measured in terms of tongue height (Z3-Z1) to determine whether the productions differed when producing /i:/ in Arabic and English. Figure 3.24 shows Z3-Z1 of the Arabic and English /i:/ produced by A mono, E mono and A-E and E-A bilinguals. To determine whether tongue height differed between the monolingual groups, I compared the Z3-Z1 of A mono and E mono speakers. As shown in Figure 3.24, the Z3-Z1 was numerically higher for E mono than for A mono, suggesting that E mono speakers are likely to produce the /i:/ vowel with a higher tongue position than are A mono speakers.

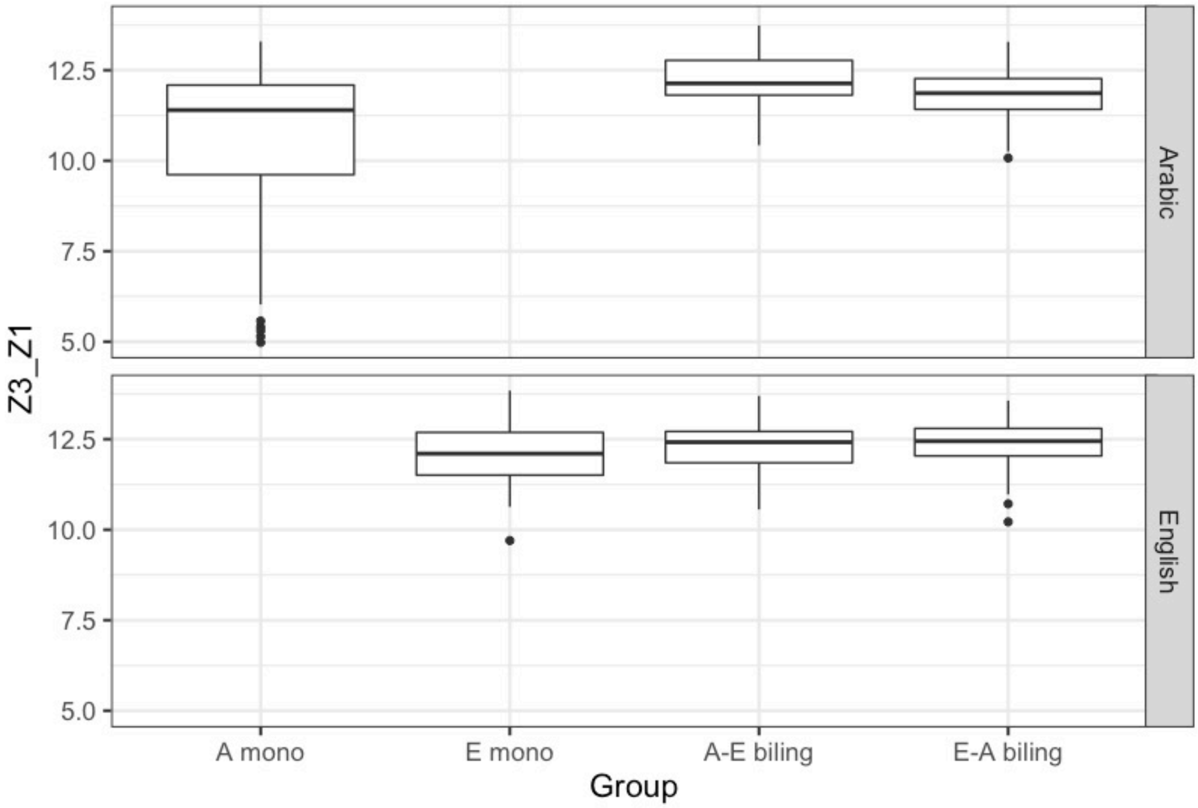


Figure 3.24. Boxplots for the Z3-Z1values (Bark) of the Arabic and English /i:/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

I employed linear mixed effects models to determine whether the monolingual Arabic and English speakers shown in Figure 3.24 differ statistically significantly in terms of the tongue height of /i:/. Table 3.35 shows that the tongue height was statistically significantly higher for E mono than for A mono speakers in the production of /i:/. Moreover, a significant difference was evident in the speaking condition, with a significantly higher tongue position in the sentence compared to in the isolation speaking condition. The speaker group by speaking condition interaction was not significant.

Table 3.35. Results of the Z3-Z1Comparison between A Mono and E Mono in Producing /i:/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.235	0.309	36.350	33.064	< .0001
Speaker group	2.176	0.289	167.580	7.522	< .0001
Speaking condition	0.744	0.272	158.110	2.736	< .001
Speaker group x speaking condition	-0.732	0.384	-157.890	1.907	= .058

Bilinguals

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z1 in Producing /i:/ in Arabic

As the Z3-Z1of the /i:/ vowel of the monolingual participants differed, I compared the bilinguals' productions with those of the monolinguals. First, I compared the Arabic bilinguals' productions with the A mono speakers' productions. Figure 3.24 shows the different Z3-Z1 values of the A mono speakers, E-A bilinguals and A-E bilinguals when producing the /i:/ vowel in Arabic. To determine whether the A-E and E-A bilinguals differed statistically significantly from A mono in terms of tongue height in producing the Arabic /i:/, linear mixed effects models were used. The results revealed that both bilingual groups differed statistically significantly from A mono in their Z3-Z1 when producing the Arabic /i:/. Thus, the A-E bilinguals showed attrition in their L1 when producing the Arabic /i:/, and the E-A bilinguals did not achieve L2 target-like productions in terms of Z3-Z1when producing /i:/ in Arabic. Both bilingual groups' Arabic productions were moving towards the English monolingual norm. In addition, speaking condition had a significant effect, showing that all participants produced the /i:/ vowel with a significantly higher tongue height in the sentence condition than in the isolation condition.

Table 3.36. Results of the Z3-Z1Comparison between A Mono and A-E and E-A Bilinguals in Producing /i:/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	10.232	0.224	81.490	45.527	< .0001
A mono vs A-E biling	2.184	0.275	251.010	7.927	< .0001
A mono vs E-A biling	1.759	0.270	250.090	6.512	< .0001
Speaking condition	0.791	0.268	250.360	2.941	= .003
A mono vs A-E biling x speaking condition	-1.132	0.382	251.840	-2.963	= .003
A mono vs E-A biling x speaking condition	-1.161	0.380	250.180	-3.056	= .002

Since significant differences were evident between the A-E and E-A bilinguals and the Arabic monolinguals, I considered the language use and phonetic aptitude factors, and I used linear mixed effects models to determine whether aptitude and language use influenced the Z3-Z1 for /i:/ in Arabic produced by the A-E and E-A bilinguals. The analysis shows that there was no main effect of aptitude, no main effect of language use and no interaction between the two.

Table 3.37. Z3-Z1 Comparison between A-E bilingual with Low and High Aptitude and More Arabic and More English in Producing /i:/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	388.496	39.176	11	9.917	< .0001
Aptitude	-16.543	45.237	11	-0.366	= .722
Language use	11.149	47.981	11	0.232	= .821
Aptitude x language use	-5.853	61.943	11	-0.094	= .926

Table 3.38. Z3-Z1 Comparison between E-A Bilingual with Low and High Aptitude and More Arabic and More English in Producing /iː/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	420.11	24.63	11	17.058	< .0001
Aptitude	-36.37	37.62	11	0.967	= .354
Language use	-45.73	34.83	11	-1.313	= .216
Aptitude x language use	84.97	51.27	11	1.657	= .126

Since the results in Table 3.38 also show significant interactions with speaking condition, I used linear mixed effects models to perform separate analyses for each group of bilinguals. Table 3.39 demonstrates that no statistically significant differences exist for A-E bilinguals for speaking condition in producing /iː/, suggesting that the A-E bilinguals produce the Arabic /iː/ vowel similarly in a carrier phrase and in isolation. By contrast, Table 3.40 shows that the E-A bilingual produce the Arabic /iː/ vowel with a lower tongue position in isolation than in a carrier phrase.

Table 3.39. A-E Bilinguals Producing Arabic /iː/ and Speaking Condition

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	375.65	13.572	16.499	27.68	< .0001
A-E biling x speaking condition	13.05	7.193	74.71	1.814	= .0737

Table 3.40. E-A Bilinguals Producing Arabic /iː/ and Speaking Condition

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	389.854	13.173	16.25	29.64	< .0001
E-A biling x speaking condition	23.104	7.083	74	3.262	< .001

Comparison of E-A and A-E Bilinguals with E Monolingual in the Z3-Z1 in Producing /iː/ in English

Since the Z3-Z1 of the /iː/ vowel differed across the monolingual speakers, I compared the English productions of the bilinguals with those of the English monolinguals. Figure 3.24 shows rather similar Z3-Z1 values for E mono and A-E and E-A bilinguals when producing the /iː/ vowel in English. The linear mixed effects models showed that both bilingual groups differed significantly from E mono speakers in that they produced the English /iː/ vowel with significantly higher tongue position than E mono. As a result, the E-A bilinguals show attrition in their L1 and the A-E bilinguals did not produce the /iː/ like English monolinguals in terms of tongue height. Furthermore, speaking condition and the speaker group by speaking condition interactions revealed no significant differences.

Table 3.41. Z3-Z1Comparison between E Mono and A-E and E-A Bilinguals in Producing the English /iː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	12.110	0.120	66.060	100.782	< . 0001
E mono vs A-E biling	0.285	0.143	259.910	1.986	= . 048
E mono vs E-A biling	0.397	0.1444	259.980	2.755	= . 006
Speaking condition	0.011	0.138	239.830	0.084	= .933
E mono vs A-E biling x speaking condition	-0.077	0.196	239.830	-0.397	= .691
E mono vs E-A biling x speaking condition	-0.177	0.196	239.880	-0.902	= .367

Since significant differences were evident between the A-E and E-A bilinguals and the English monolinguals, I considered the language use and phonetic aptitude factors, and I used linear mixed effects models to determine whether aptitude and language use influenced the Z3-Z1 for /iː/ in English produced by the A-E and E-A bilinguals. The analysis shows that there was no main effect of aptitude and language use among the A-E and E-A bilinguals, no interaction between the aptitude and language use.

Table 3.42. Z3-Z1 Comparison between A-E bilingual with Low and High Aptitude and More Arabic and More English in Producing /iː/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	12.478	0.423	11	29.448	< .0001
Aptitude	0.329444	0.489	11	-0.673	= .515
Language use	0.114	0.518	11	0.221	= .829
Aptitude x language use	0.001	0.669	11	-0.002	= .998

Table 3.43. Z3-Z1 Comparison between E-A Bilingual with Low and High Aptitude and More Arabic and More English in Producing /iː/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	12.1126	0.204	11	159.233	< .0001
Aptitude	0.232	0.311	11	0.746	= .471
Language use	0.647	0.288	11	2.247	= .646
Aptitude x language use	-0.630	0.423	10.977	1.489	= .165

Results for the Z3-Z2 Values of the /iː/ Vowel in Arabic and English

Comparison of Monolingual Patterns

Figure 3.25 shows the Z3-Z2 values of the Arabic and English /iː/ produced by A mono, E mono and the A-E and E-A bilinguals. Figure 3.25 shows numerically higher Z3-Z2 values for A mono than for E mono speakers, indicating that A mono tend to produce the /iː/ vowel with a less frontal tongue position than do E mono speakers.

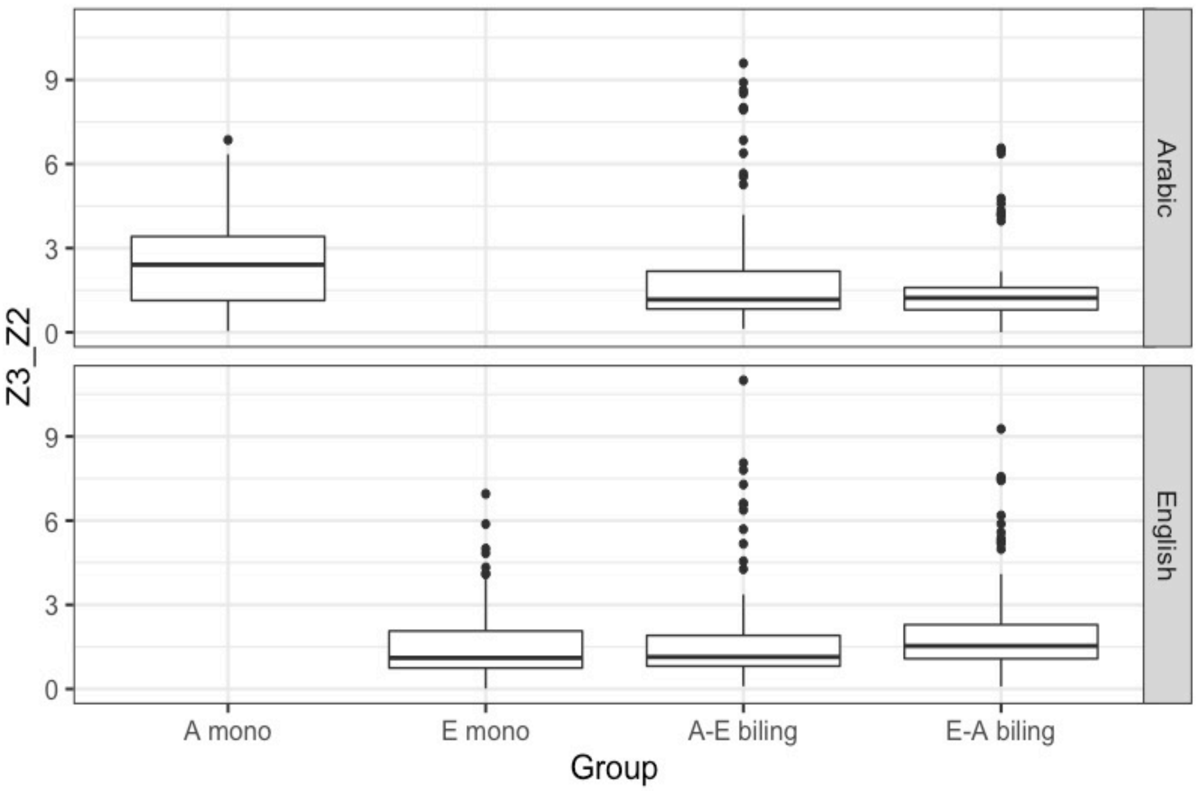


Figure 3.25. Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /iː/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals

Linear mixed effects models were used to determine whether the monolingual Arabic and English speakers differ statistically significantly in terms of backness/frontedness (Z3-Z2) when producing /iː/. Table 3.44 shows a statistically significant difference in the Z3-Z2 between A mono and E mono speakers when producing /iː/. A mono speakers produced the /iː/ in a significantly less frontal position than did the E mono speakers. No significant differences were evident in the speaking condition or in the speaker group by speaking condition interaction.

Table 3.44. Z3-Z2 Comparison of A Mono and E Mono Speakers in Producing /iː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	2.572	0.246	51.860	10.434	< .0001
Speaker group	-1.231	0.282	171.690	-4.353	< .0001
Speaking condition	-0.348	0.270	156.410	-1.289	= .199
Speaker group x speaking condition	0.597	0.381	155.980	1.565	= .120

Bilingals

Comparison of E-A and A-E bilinguals with A Monolingual in the Z3-Z2 in producing /i:/ in Arabic

Because the monolingual participants differed in their productions of the Arabic /i:/ vowel in terms of the Z3-Z2, I compared the bilinguals' productions with those of the Arabic monolinguals. As shown in Figure 3.25, A mono and A-E and E-A bilinguals differed numerically in the production of /i:/ in Arabic with both bilingual groups producing /i:/ in a more fronted position. Linear mixed effects models were used to determine whether the E-A and A-E bilinguals displayed statistically significant differences in terms of the Z3-Z2 when producing the Arabic /i:/. The results in Table 3.45 revealed that the E-A bilingual group were significantly different from A mono speakers. While the A-E bilingual group did not differ from the A mono. Thus, the A-E bilingual group have native-like productions in terms of tongue frontedness when producing the Arabic /i:/, and the E-A bilingual group did not approximate the native Arabic target when producing /i:/.

Table 3.45. Z3-Z2 Comparison of A Mono and A-E and E-A Bilinguals in Producing /i:/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	2.571	0.274	137.090	9.375	< .0001
A mono vs A-E biling	-0.681	0.367	251.680	-1.852	= .065
A mono vs E-A biling	-1.335	0.360	250.110	-3.703	< .0001

As significant differences exist between the E-A bilinguals and the monolingual Arabic speakers, I examined the language use and phonetic aptitude factors, using linear mixed effects models to determine whether aptitude and language use influenced the Z3-Z2 for E-A bilinguals' Arabic /i:/. The analyses in Tables 3.46 showed no main effects of aptitude, no main effects of language use and no interactions between the two.

Table 3.46. Z3-Z2 Comparison of E-A Bilinguals with Low and High Aptitude and More Arabic and More English in Producing /i:/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	2332.8	169.5	11	13.77	< .0001

Aptitude	337.1	258.9	11	1.302	= .22
Language use	111.4	239.7	11	0.47	= .651
Aptitude x language use	-314.8	352.8	11	-0.892	= .391

Comparison of A-E and E-A Bilinguals with E Monolingual in the Z3-Z2 in Producing /iː/ in English

Since the Z3-Z2 values for the /iː/ vowel differed between A mono and E mono, I also explored how bilingual English productions compared with those of E mono speakers. Figure 3.17 shows similar Z3-Z2 values for the English /iː/ for E mono speakers and the E-A bilinguals and A-E bilinguals. I used linear mixed effects models to determine whether the E mono, E-A bilinguals and A-E bilinguals differed statistically significantly in terms of tongue backness/frontedness (Z3-Z2) when producing the English /iː/. The results in Table 3.47 reveal that E-A and A-E bilinguals did not differ significantly from E mono speakers in this regard, indicating that the E-A bilinguals showed no L1 attrition and that A-E bilinguals achieved target-like productions when producing the English /iː/.

Table 3.47. Z3-Z2 Comparison of E Mono and A-E and E-A Bilinguals in Producing /iː/ in English

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	1.564	0.317	57.540	4.932	< .0001
E mono vs A-E biling	0.248	0.345	258.520	0.718	= .453
E mono vs E-A biling	0.539	0.347	258.560	1.551	= .122

3.6.6 Results for the Z3-Z1 Values of the /uː/ Vowel in Arabic and English Across Speaker Groups

The final vowel I consider is the long vowel /uː/. Figure 3.26 shows the mean Z3-Z1and Z3-Z2 values (in Bark) of the long vowel /uː/ for Arabic monolinguals, English monolinguals, A-E bilinguals and E-A bilinguals. Z3-Z1 represents the vertical dimension and Z3-Z2 represents the horizontal dimension. As with the /iː/ sound, every group has a small range when it comes to Z3-Z1, except for the Arabic monolinguals who have a greater range due to some outliers.

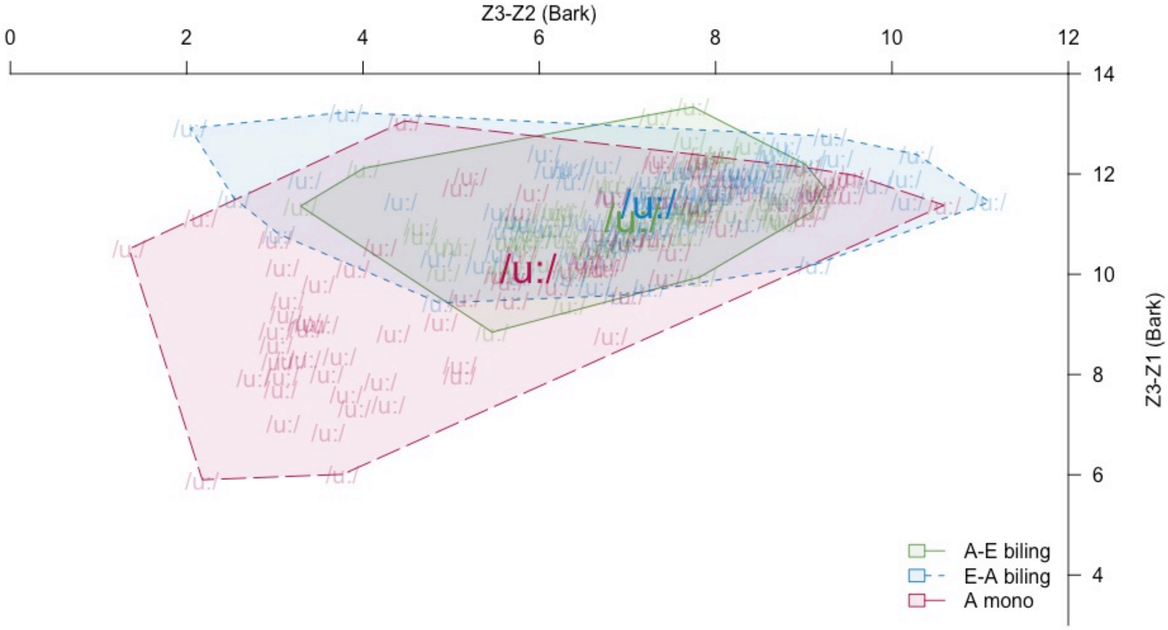


Figure 3.26. Mean Z3-Z1and Z3-Z2 for the Arabic /uː/ vowel produced by Amono, A-E biling and E-A biling

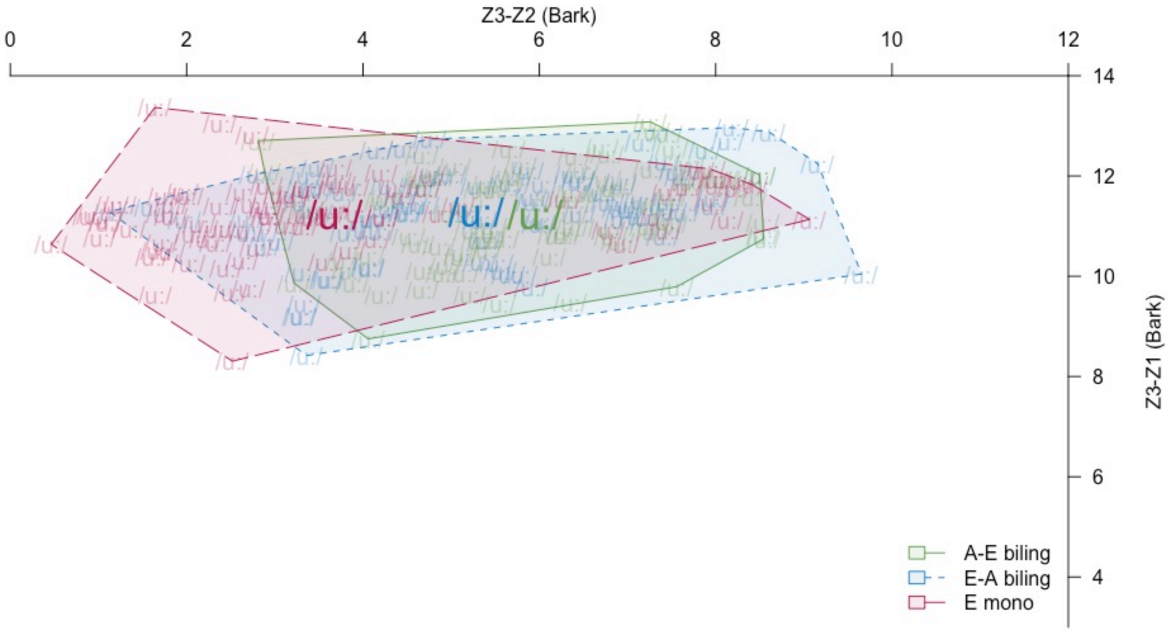


Figure 3.27. Mean Z3-Z1and Z3-Z2 for the English /u:/ vowel produced by E mono, A-E biling and E-A biling

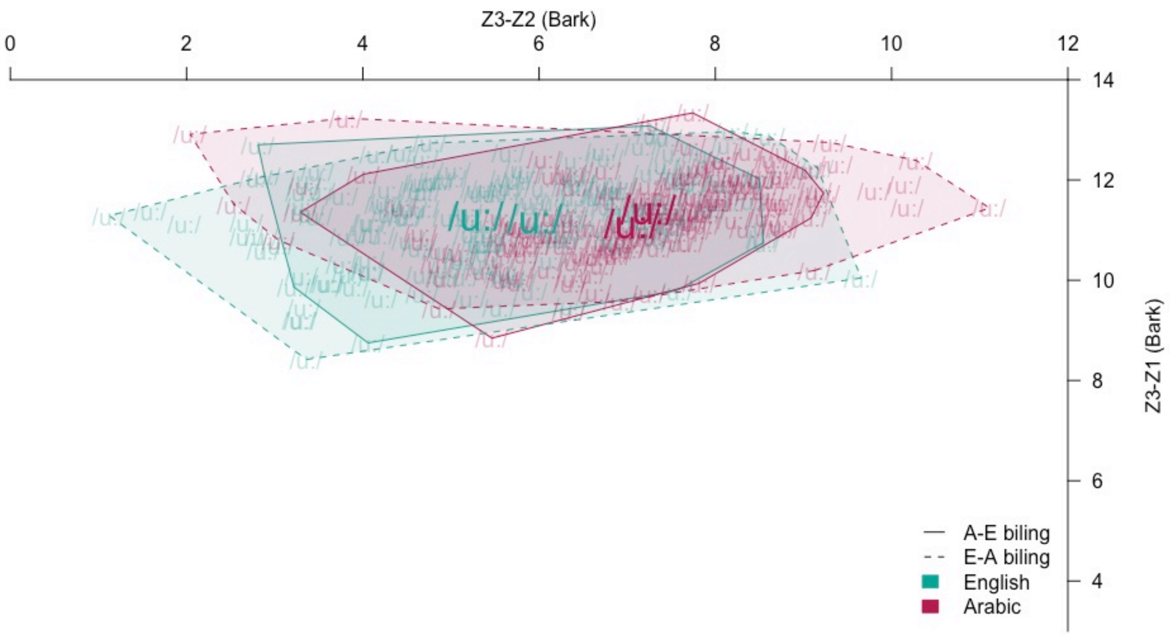


Figure 3.28. Mean Z3-Z1and Z3-Z2 for /u:/ vowel produced by, A-E biling and E-A biling in Arabic and English

Comparison of Monolingual Patterns

Again, I first determined whether A mono and E mono speakers varied in their productions of the /u:/ vowel in terms of tongue height (Z3-Z1). Figure 3.29 shows the Z3-Z1 values of the Arabic and English /u:/ for A mono and E mono speakers and A-E and E-A bilinguals. To determine whether the tongue height when producing the /u:/ vowel differed between the two languages, I compared the Z3-Z1 of the A mono and E mono speakers. Figure 3.29 shows that the Z3-Z1 of E mono was numerically higher than A mono, suggesting that E mono speakers tend to produce the /u:/ vowel with a higher tongue position than A mono speakers.

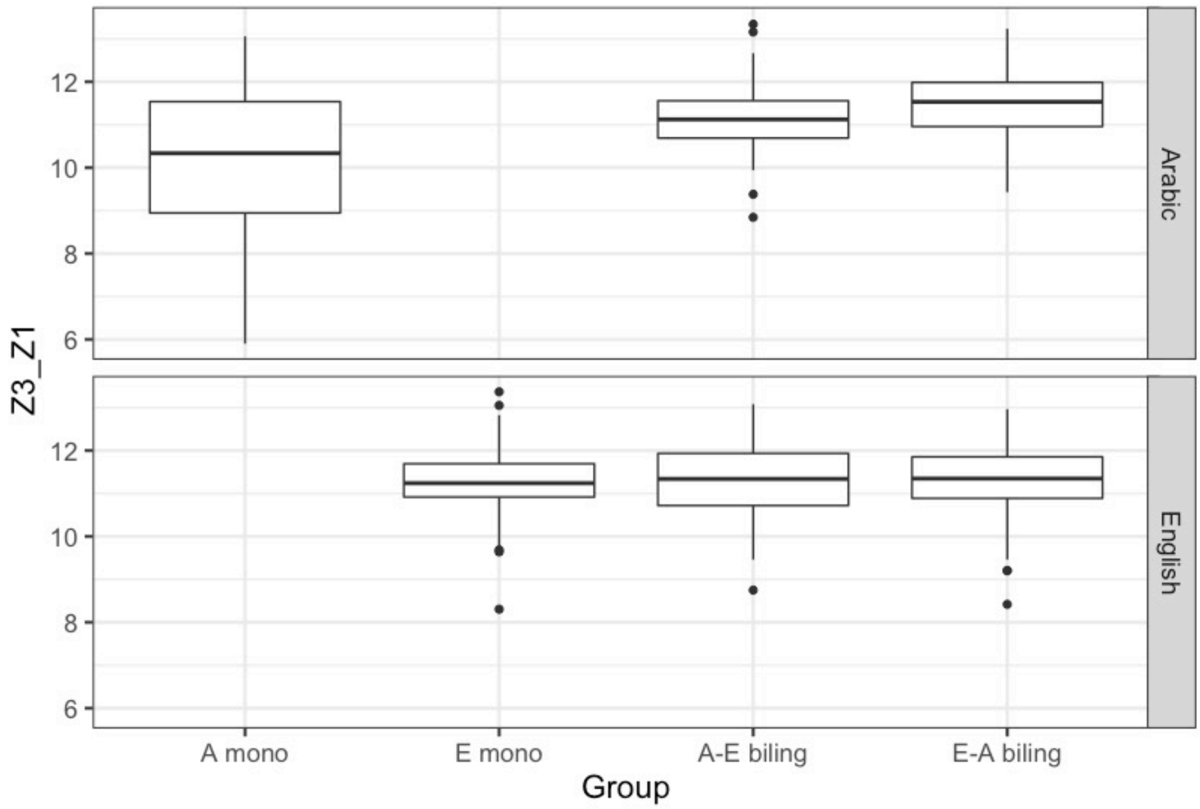


Figure 3.29. Boxplots for the Z3-Z1values (Bark) of the Arabic and English /uː/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

I used linear mixed effects models to determine (1) whether E mono and A mono speakers differed statistically significantly when producing /uː/ in terms of tongue height and (2) whether the productions differed in a carrier phrase compared to in isolation. Table 3.48 shows the results. The analysis showed that speaker group did not differ significantly, showing similar Z3-Z1 values for A mono and E mono speakers. Moreover, speaking condition had no significant effect, and the speaker group by speaking condition interaction was not significant.

Table 3.48. Z3-Z1Comparison of A Mono and E Mono in Producing /uː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	11.302	0.140	55.950	80.550	< .0001
Speaker group	0.104	0.133	261.810	0.787	= .432
Speaking condition	0.126	0.144	246.310	-0.879	= .380
Speaker group x speaking condition	0.033	0.176	246.400	0.189	= .851

Since monolingual participants did not differ in the Z3-Z1 of their /u:/ productions, I did not compare the bilinguals' productions of the /u:/ vowel in terms of Z3-Z1 with those of the monolinguals.

Results for the Z3-Z2 Values of the /u:/ Vowel in Arabic and English

Comparison of Monolingual Patterns

Finally, I determined whether A mono and E mono speakers differed in their productions of the /u:/ vowel in terms of the Z3-Z2, that is, in terms of the degree of frontedness/backness. Figure 3.30 shows the Z3-Z2 values of the Arabic and English /u:/ for the E-A and A-E bilinguals and the A mono and E mono speakers. I first compared the Z3-Z2 of both monolingual groups to determine whether the Z3-Z2 of the /u:/ sounds differed between the two groups. Figure 3.30 suggests that the Z3-Z2 for E mono is lower than that for A mono, suggesting that A mono speakers are likely to produce the /u:/ with less fronting than are E mono speakers.

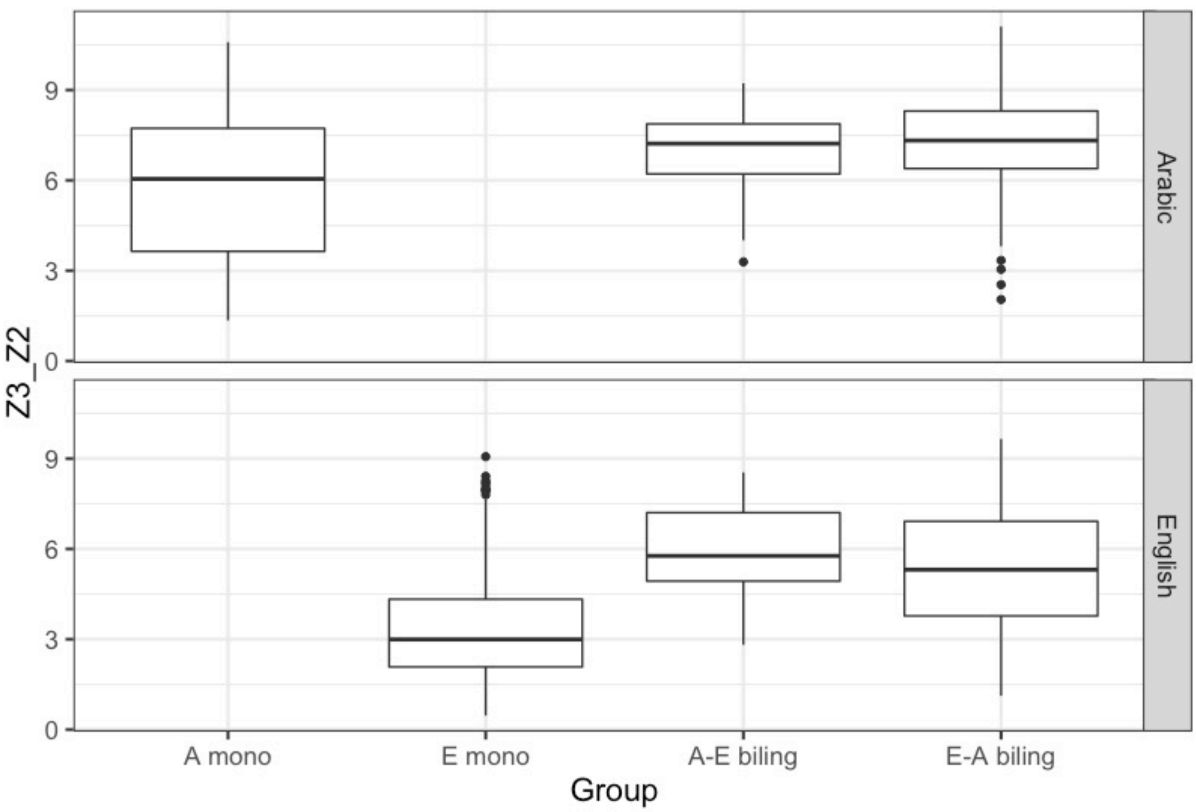


Figure 3.30. Boxplots for the Z3-Z2 values (Bark) of the Arabic and English /u:/ produced by A mono, E mono, A-E bilinguals and E-A bilinguals.

Linear mixed effects models were used to determine whether the monolingual Arabic and English speakers shown in Figure 3.30 differed statistically significantly in terms of the Z3-Z2 values for the long vowel /uː/. Table 3.49 shows a statistically significant difference in the Z3-Z2 between A mono and E mono speakers. The A mono group produced the /uː/ in a less fronted position. No differences were evident in the speaking condition or in the speaker group by speaking condition interaction.

Table 3.49. Z3-Z2 Comparison of A Mono and E Mono in Producing /uː/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	3.611	0.386	44.500	9.338	< .0001
Speaker group	2.459	0.365	261.020	6.732	< .0001
Speaking condition	0.016	0.395	241.380	0.041	= .967
Speaker group x speaking condition	0.449	0.485	241.490	0.925	= .356

Comparison of E-A and A-E Bilinguals with A Monolingual in the Z3-Z2 in Producing /uː/ in Arabic

Because the monolingual participants differed in the Z3-Z2 values for the /uː/ vowel, I examined how the productions of bilinguals compared with those of monolinguals. I first compared A mono productions with the Arabic productions of bilinguals using linear mixed effect smodels.

As statistically significant differences exist between the A-E bilingual and the monolingual Arabic speakers, I consider the language use and phonetic aptitude factors. To determine whether aptitude and language use influenced the Z3-Z2 values for the Arabic /uː/ for A-E bilinguals, I used linear mixed effects models. The statistical analysis in Table 3.50 shows that there is no main effect of aptitude, no main effect of language use and no interaction between the two.

Table 3.50. Z3-Z2 Comparison of A-E Bilinguals with Low and High Aptitude and More Arabic and More English in producing /u:/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	451.47	32.53	11.00	13.880	< .0001
Aptitude	-36.55	37.56	11.00	-0.973	= .351
Language use	-18.14	39.84	11.00	-0.455	= .658
Aptitude x language use	-10.32	51.43	11.00	-0.201	= .845

Comparison of A-E and E-A Bilinguals with E Monolingual in the Z3-Z2 in Producing /u:/ in English

Because the Z3-Z2 values of the /u:/ vowel differed between the A mono and E mono groups, I next considered how bilinguals' English productions compared with the E mono productions. I used linear mixed effects models to determine whether E mono, E-A bilinguals and A-E bilinguals differed statistically significantly in terms of tongue backness/frontedness when producing the English /u:/. The results in Table 3.51 reveal that the E-A and A-E bilinguals differed statistically significantly from E mono speakers, meaning that the E-A bilinguals showed L1 attrition and the A-E bilinguals failed to achieve target-like productions in terms of frontedness when producing the English /u:/. Both groups of bilinguals moved towards the Arabic norm in terms of Z3-Z2 values for the English /u:/.

Table 3.51. Z3-Z2 Comparison of E Mono and A-E and E-A Bilinguals in Producing the English /u:/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	3.626	0.346	42.550	10.457	< .0001
E mono vs A-E biling	1.690	0.374	259.570	4.515	< .0001
E mono vs E-A biling	2.480	0.371	259.810	6.685	< .0001

Since significant differences were evident between the A-E and E-A bilinguals and the monolingual English speakers, I considered the language use and phonetic aptitude factors, using linear mixed effects models to determine whether aptitude and language use influenced the Z3-Z2 values for the English /u:/ for the A-E and E-A bilinguals. Tables 3.52

and 3.53 show the statistical analyses, which indicate that no main effects of aptitude or language use exist and there is no interaction between the two.

Table 3.52. Z3-Z2 Comparison of A-E Bilinguals with Low and High Aptitude and More Arabic and More English in Producing /u:/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	1319.6	129.81	10.99	10.17	< .0001
Aptitude	-169.1	150	11.02	-1.13	= .284
Language use	-194.61	158.6	10.9	-1.23	= .25
Aptitude x language use	190.62	204.99	10.93	0.930	= .373

Table 3.53. Z3-Z2 Comparison of E-A Bilinguals with Low and High Aptitude and More Arabic and More English in Producing /u:/ in English

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	1259.30	180.5	11	6.98	< .0001
Aptitude	181.8	275.7	11	0.66	= .523
Language use	45.92	255.3	11	0.18	= .861
Aptitude x language use	-224.3	375.72	11	-0.597	= .563

3.7 Discussion

This chapter investigated L2 vowel acquisition and L1 vowel attrition by measuring formant values in the productions of shared vowels in late E-A and A-E bilinguals who speak two typologically different languages. This chapter also explored whether bilinguals' phonetic aptitude and/or language use influences how target-like their productions of vowels in their L1 and L2 were. Table 3.54 summarises the results of this investigation. Since there were no effects of aptitude or language use for any of the vowels, these two factors are not considered in the table.

Table 3.54. Schematic Summary of the Results Concerning Vowel Production among Arabic and English Monolinguals and Bilinguals

Vowel	Monolinguals	Bilinguals vs A mono	Bilinguals vs E mono
/a/	Z3-Z1 (height) • E mono < A mono • sentence > isolation Z3-Z2 (frontedness) • E mono > A mono • sentence = isolation	Z3-Z1 (height) • E-A biling = A mono • A-E biling = A mono • sentence > isolation Z3-Z2 (frontedness) • E-A biling = A mono • A-E biling = A mono	Z3-Z1 (height) • E-A biling = E mono • A-E biling = E mono • sentence = isolation Z3-Z2 (frontedness) • E-A biling = E mono • A-E biling = E mono
/ɪ/	Z3-Z1 (height) • E mono > A mono • sentence = isolation Z3-Z2 (frontedness) • E mono < A mono • sentence = isolation	Z3-Z1 (height) • A-E biling > A mono • E-A biling > A mono • sentence > isolation Z3-Z2 (frontedness) • E-A biling =A mono • A-E biling =A mono	Z3-Z1 (height) • E-A biling = E mono • A-E biling = E mono • sentence = isolation Z3-Z2 (frontedness) • E-A biling > E mono • A-E biling > E mono
/ʊ/	Z3-Z1 (height) • E mono > A mono • sentence = isolation Z3-Z2 (frontedness) • E mono = A mono • sentence = isolation	Z3-Z1 (height) • A-E biling > A mono • E-A biling > A mono Z3-Z2 (frontedness) • N/A	Z3-Z1 (height) • E-A biling = E mono • A-E biling = E mono Z3-Z2 (frontedness) • N/A
/ɑː/	Z3-Z1 (height) • E mono = A mono • sentence = isolation Z3-Z2 (frontedness) • E mono > A mono • sentence = isolation	Z3-Z1 (height) • N/A Z3-Z2 (frontedness) • E-A biling = A mono • A-E biling = A mono	Z3-Z1 (height) • N/A Z3-Z2 (frontedness) • E-A biling = E mono • A-E biling = E mono
/iː/	Z3-Z1 (height) • E mono > A mono • isolation < sentence • speaker group by	Z3-Z1 (height) • E-A biling > A mono • A-E biling > A mono • sentence > isolation	Z3-Z1 (height) • E-A biling < E mono • A-E biling < E mono • sentence = isolation

	speaking condition interaction Z3-Z2 (frontedness) • E mono > A mono • sentence = isolation	• speaker group by speaking condition interaction Z3-Z2 (frontedness) • A mono > E-A biling • A mono = A-E biling	Z3-Z2 (frontedness) • E-A biling > E mono • A-E biling > E mono
/uː/	Z3-Z1 (height) • E mono = A mono • sentence = isolation Z3-Z2 (frontedness) • E mono > A mono • sentence = isolation	Z3-Z1 (height) • N/A Z3-Z2 (frontedness) • A mono = E-A biling • A mono < A-E biling	Z3-Z1 (height) • N/A Z3-Z2 (frontedness) • E mono < E-A biling • E mono < A-E biling

Note: N/A= no analyses for the bilinguals as the monolingual participants did not differ.

This chapter tests five hypotheses. Hypothesis 1 comprises three aspects: I expected that (1) native Arabic speakers would tend to produce /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position compared to native English speakers (Hypotheses 1a); (2) native Arabic speakers would tend to produce /ɑː/ and /a/ with a higher tongue position (Z3-Z1) compared to native English speakers (Hypothesis 1b); and (3) native Arabic speakers would tend to produce /ɪ/, /iː/, /uː/, /ʊ/, /a/ and /ɑː/ with a less frontal tongue position (Z3-Z2) than do native English speakers (Hypothesis 1c). A direct comparison of the Z3-Z1 and Z3-Z2 of each vowel separately produced by monolinguals partially confirmed Hypothesis 1: All vowels differed significantly across Arabic and English monolinguals in terms of tongue height, except for /ʊ/ and /ɑː/, which were similar in this respect in both monolingual groups. In addition, most – but not all – vowels differed in the predicted direction across the monolingual groups. In details, Arabic speakers produced /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position compared to native English speakers which confirmed Hypothesis 1a. Moreover, monolingual Arabic speakers produced /a/ with higher tongue position and produced /ɑː/ similarly in terms of tongue height compared to monolingual English speakers, which partially confirmed Hypothesis 1b. In terms of frontedness, the Arabic monolinguals produced /a/, /ɪ/, /iː/ and /uː/ with a less frontal tongue position (Z3-Z2) than did native English speakers, while /ɑː/ in Arabic had numerically lower Z3-Z2 values and was thus more fronted than English /ɑː/,

which also partially confirmed Hypothesis 1c. In addition, both monolingual groups produced the /ʊ/ with similar tongue frontedness which did not confirme Hypotheses 1c.

I expected that the bilingual participants would pronounce L2 vowels as they do in their L1. Thus, bilinguals' L2 vowels were expected to differ from those of monolingual participants (Hypothesis 2). Comparing the Z3-Z1 and Z3-Z2 of bilinguals' L2 with those of monolinguals revealed that only the E-A bilinguals' vowel productions differed from those of Arabic monolinguals in terms of tongue height and tongue frontedness; A-E bilinguals produced some of the shared English vowels similarly to native monolinguals, for example, /ɑː/ and /a/. Thus, Hypothesis 2 was partially supported.

Moreover, I expected that, similar to other elements of their native languages, bilinguals would exhibit attrition when producing vowels (Hypothesis 3). While I identified attrition in the A-E bilinguals' productions of the Arabic /i/ and /iː/ vowels, the E-A bilinguals exhibited attrition for the vowels /uː/ and /iː/. They produced the English /uː/ with a less fronted tongue position and the English /iː/ with a lower tongue position than the English monolinguals.

I predicted that language use would influence L1 attrition and L2 acquisition (Hypothesis 4). However, there was no evidence suggesting that this was the case and, therefore, the present results did not confirm Hypothesis 4. Last, based on Mennen's (2004) research, in which participants had native-like productions either in both languages or in neither, I tentatively assumed that participants with high phonetic aptitude would show native-like production of *shared vowels* in their L1 and L2 (Hypothesis 5). This hypothesis was not confirmed because, similar to language use, phonetic aptitude did not influence L1 attrition and L2 acquisition. I discuss these results in greater detail in the following sections.

3.7.1 Monolinguals

The results presented in this chapter confirm Hypothesis 1a, which proposed that native Arabic speakers tend to produce /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position (Z3-Z1) than do native English speakers. According to the current data, Arabic monolinguals did indeed produce /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position, confirming this Hypothesis. Hypothesis 1b, which suggested that native Arabic speakers tend to produce /ɑː/ and /a/ with a higher tongue position than native English speakers. Arabic monolinguals produced the short vowel /a/ with a higher tongue position (higher Z3-Z1) than did the English monolinguals, which partially confirms Hypothesis 1b; however, in contrast with the previous literature, A mono and E mono produced the /ɑː/ in a similar way. Finally, Hypothesis 1c, which proposed that native Arabic speakers produced all six vowels – /ɪ/, /iː/, /ʊ/, /uː/, /a/, and /ɑː/ – with a less fronted tongue position (Z3-Z2) than did native English speakers, was also partially confirmed. According to the current data, the monolinguals' Arabic production of /ɑː/ in Arabic have numerically lower Z3-Z2 values and were thus more fronted compared to their English counterparts, which conflicts with Hypothesis 1c, while the Arabic monolinguals produced /a/, /ɪ/, /iː/ /uː/ with a higher Z3-Z2 value than did English monolinguals, which supports Hypothesis 1c. In addition, the present study reports that /ʊ/ has comparable Z3-Z2 values in English and Arabic, which did not support Hypothesis 1c.

Thus, the current findings concerning Arabic monolinguals only partially confirm the suggestions of previous studies. In particular, Alotaibi and Hussain (2009) found that native Arabic speakers tend to produce /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position and /aː/ and /a/ with a higher tongue position (Z3-Z1) than do native English speakers. They also found that Arabic monolinguals tend to produce /iː/, /ɪ/, /uː/, /ʊ/, /ɑː/ and /a/ with a less frontal tongue position (Z3-Z2). There are several reasons that potentially explain the discrepancy between the current results and the findings of Alotaibi and Hussain (2009). In the present chapter, I collected vowel data from male and female native Arabic speakers, whereas Alotaibi and Hussain (2009) collected data from nine male speakers and one child, which may have affected the results. Specifically, children tend to have a very high fundamental frequency, which makes it difficult to estimate formant frequencies (Story & Bunton, 2015). Moreover, one of the male participants in Alotaibi and Hussain's (2009) study was from Egypt, which may have also affected their results, as the Egyptian dialect of Arabic is quite different from MSA. In addition, the present chapter used the /hVd/ frame, and the target vowels were embedded in this frame to control for the effect of phonetic context. By contrast, Alotaibi and Hussain (2009) used different consonants in the same CVC frame.

The current results concerning native English speakers also only partially align with those of Deterding (2006). There are several possible reasons for the discrepancy between these results. I used a controlled environment for vowels in that I used the /hVd/ frame to control for the effect of phonetic context, while Deterding (2006) used the MARSEC database, which contains a set of monologues, including news articles, broadcasts and commentary, produced by the BBC. In addition, Deterding (2006) obtained vowel data from equal genders; five males and five females, while the present data were collected from two males and 13 females.

3.7.2 Bilinguals: L2 Acquisition

I discuss the data obtained from the bilingual participants from an L2 acquisition perspective first, followed by an L1 attrition perspective. The data in this chapter partially confirm Hypothesis 2. E-A bilinguals did not form a native-like new category for the L2 vowels /ʊ/, /u:/, /ɪ/ and /i:/, while they formed a native-like new category for the L2 vowels /ɑ:/, /a/. In contrast, the A-E bilinguals formed a native-like new category for the L2 vowels /ɑ:/, /a/, and /ʊ/ and did not form a native-like new category for /ɪ/, /i:/, and /u:/. Thus, some results from the E-A and A-E bilinguals align with those of previous studies supporting the SLM and PAM-L2 theories (e.g. Munro et al., 1996; Flege, 1987; Flege & Hillenbrand, 1984). In more detail, the E-A bilinguals are native-like in their L2 Arabic for the vowels /ɑ:/, /a/ and /u:/, whereas the A-E bilinguals are native-like in their L2 English for the vowels /ɑ:/, /a/, and /ʊ/. Thus, both E-A and E-A bilinguals produced the /a/, /ɑ:/ vowels like a native in both languages, suggesting that they formed a separate, native-like category for the Arabic and Engish /a/ and /ɑ:/ in their L2.

Overall, the L2 results for the E-A and A-E bilinguals are compatible with the SLM and PAM-L2 theories. The comparison of the vowel productions by the two monolingual participant groups suggests that the vowels analysed here are indeed similar vowels across the two languages, in the sense that they reveal significant differences in terms of tongue height and/or frontedness across the two languages, but occupy overlapping ranges in vowel space. The SLM and PAM-L2 would therefore predict that such similar, but not identical sounds, are difficult, but not impossible, to learn. In line with these predictions, the highly advanced L2 learners in the current study, who have lived in the L2 environment for a long time, still are not native-like in their L2 productions of most of these vowels. This suggests that these similar, but not identical, sounds are indeed difficult to learn. In contrast, the current results are less compatible with the L2LP model, which suggests that *dis*similar sounds post particular difficulties. According to the L2LP model, the similar sounds analysed in the current study should therefore have been relatively easy to learn, and one should have expected native-like L2 productions for these similar vowels in such a highly proficient learner group as in the current study. It needs to be mentioned though that I only analysed similar vowels, but not dissimilar vowels, in the current study. And while the results from the similar vowels are less compatible with the L2LP model, and more compatible with the SLM and PAM-L2, it needs to be mentioned that stronger evidence for or against these models would have come from an analysis of both similar and dissimilar vowels. Future studies should therefore expand on the analyses conducted here, and also analyse dissimilar vowels.

The results for A-E bilinguals who formed a native-like new category for the L2 vowels /ɑː/, /a/, and /ʊ/ and E-A bilinguals who formed a native-like new category for the L2 vowels /ɑː/, /a/ align with those of Wu and Shih (2012), which – like the current study – did not completely support the predictions of the SLM theory. The current results differ from those of Munro et al. (1996), who showed that early bilinguals who learned the L2 during childhood were able to perceive and produce the vowels of the L2 more accurately than late bilinguals who learned English after the onset of puberty. In contrast, the current study finds native-like vowel productions in late bilinguals for some vowels. There are several possible reasons for this discrepancy. First, unlike Munro et al. (1996), I looked at highly proficient bilinguals rather than less proficient L2 learners. Even though these similar vowels were predicted to be difficult to learn, it is of course not impossible to learn them, and it seems that the highly proficient bilinguals in the current study did indeed achieve native-like pronunciations for some of the vowels. Thus, while these results go against the predictions I made, i.e. the most likely outcome, based on the SLM and PAM-L2, they are nevertheless compatible with these frameworks. Second, Munro et al. (1996) collected data from participants with different ages of L2 acquisition (e.g. before the age of 12), whereas I studied participants with an average age of 33 who had learned the L2 after puberty.

The current data concerning the acquisition of L2 vowels did not support Hypothesis 4. Both groups of A-E and E-A bilinguals, i.e. those with high and low L2 use, achieved native-like L2 pronunciation in some of the shared vowels. Overall, language use did not affect the E-A bilinguals' vowel productions, so that no group (high or low L2 language use) achieved more native-like pronunciation than the other. While the absence of a significant effect of language use does not mean that there is no effect (merely that no effect was detected), I will tentatively propose some reasons for the absence of an effect in the current data. It is possible that language use did not affect bilinguals' productions because all bilinguals had been living in the L2 environment for a long time. That is, despite differences in the amount of L2 language use, all participants had been living in an L2 language community for a long time and it is thus possible that they all had a sufficient amount of L2 language use. In other words, it is possible that the participants tested in the current study were too homogeneous in terms of language use to detect any differences across the two language use groups. Another possibility is that a certain kind of language use influences native-likeness rather than just language use. While the current study did use a detailed questionnaire to gauge both the quality and quantity of L2 language use, participants were categorized into low and high L2 language use groups based on their L2 language use as a whole, and based on a median split. It is possible that results may differ if participants were grouped based only on the quality of L2 language use, or based only on the quantity of L2 language use.

Hypothesis 5 was also not confirmed. Phonetic aptitude (low vs high aptitude) showed no significant influence on participants' vowel production; hence, there was no evidence that differences in productions between participants with high and low aptitude related to the nativeness of their productions. In other words, the present study found that neither group of bilinguals appeared more or less native-like in their L2. This finding contradicts previous research (e.g. Jilka et al., 2008; Hopp & Schmid, 2010). Jilka et al.'s (2008) work on sounds revealed that German-English bilinguals' pronunciation was influenced by the participants' phonetic aptitude. This difference in results may be because Jilka et al. (2008) had a very large sample size, whereas this study was carried out on a substantially smaller scale. Moreover, the tests used in this study to measure the aptitude are not exactly the same as those used by Jilka et al. (2008). While Jilka's tests drew on a wide range of linguistic skills, the current study attempted to select tests that were more closely focused on abilities relating to sounds.

Notably, the current chapter's results align with those of Flege et al. (1999), who found that their measure of sound processing ability contributed little to a foreign accent in learners' L2, and those of Suter (1976), who found that musical ability did not affect the degree of the learners' foreign accent.

Overall, the results from the current and previous studies suggest that some talents and skills might positively relate to better L2 pronunciation, but it is not yet clear which talents and skills are related to pronunciation and how exactly these talents and skills can be measured or tested. Further research is needed to determine which aptitude tests relate to more native-like pronunciation and which talents these tests actually measure.

3.7.3 Bilinguals: L1 Attrition

Looking at the findings from the perspective of L1 attrition, I find that both A-E and E-A bilinguals show attrition in some of the vowels of their L1, in line with Hypothesis 3. In detail, I found evidence of L1 attrition among A-E bilinguals and E-A bilinguals for the high front and high back Arabic and English vowels /uː/, /ɪ/, and /iː/, but not for the vowel /ʊ/, which only showed attrition in the A-E bilinguals, but not the E-A bilinguals. Of the vowels that showed attrition, /uː/ and /iː/ shifted towards those in the L2, most likely due to prolonged exposure. In contrast, the A-E bilinguals' Arabic productions of /ɪ/ did not move towards the monolingual English norm, but rather away from it, in a process called polarisation. This polarisation result aligns with Flege (1991) and Mayr et al. (2012). The E-A and A-E bilinguals showed no attrition in their L1 for the other vowels (the low front/back vowels /a/ and /ɑː/). These results for the A-E and E-A bilinguals align with those of previous studies (Mayr et al., 2012; Guion, 2003; Chang, 2012, 2013). Overall, in this study both bilingual groups showed attrition in almost the same vowels, namely mostly the high and low back vowels (/ɪ/, /iː/, /ʊ/ and /uː/), which across the two languages differed mostly in tongue height.

The vowel attrition data in this study did not support Hypotheses 4 and 5. In particular, the E-A and A-E bilinguals showed attrition in their L1 for some vowels, but the attrition was unaffected by language use or aptitude. To the best of my knowledge, this study is the first to examine whether aptitude and/or language use may influence the L1 attrition of vowels in highly proficient bilinguals. Again, it is difficult to speculate based on null results, but it may again be the case that the bilinguals in the current study all have sufficient exposure to the L2 so that they are relatively homogeneous in terms of their language use patterns. As mentioned before, in terms of aptitude, further studies are needed to determine which kinds of talents may influence L1 attrition and how these particular kinds of talents can be measured. It should also be noted that, even though both the L1 and L2 are flexible systems, the L1 is considered to be relatively more stable than the L2. Since language use and aptitude did not influence native-likeness in the L2, it would have been surprising to find an influence of language use or aptitude in terms of native-likeness in the L1.

3.7.4 Speaking Condition

The speaking condition (isolation and carrier phrase) was tested here to determine whether the shared vowels differ in height or frontedness when produced in isolation compared to in a carrier phrase. Speaking condition mattered only for the production of a few of the vowels. Specifically, differences were evident in the productions of /iː/ in isolation compared to in a sentence for the Arabic monolinguals, but not for the English monolinguals. In addition, speaking condition affected the production of the Arabic /iː/ in E-A bilinguals, but not in A-E bilinguals. This latter result may be due to the L2 participants hitting the L2 Arabic target more frequently in isolation than in a sentence.

3.7.5 Conclusion

To conclude, the present chapter showed that there were significant differences in the Arabic and English monolinguals' productions for all of the shared vowels. The results are also overall compatible with the predictions of the SLM and PAM-L2 theories that bilinguals' L2 vowels are likely to differ from those of monolinguals in the case of similar vowels. In contrast, the results are less compatible with the L2LP model, as such highly proficient bilinguals would be expected to have mastered the similar vowels analysed here if similar vowels were particularly easy to learn. As expected, I also found L1 attrition among the A-E bilinguals and E-A bilinguals, supporting the idea that both the L1 and L2 are flexible systems, such that L2 exposure can influence the L1. Finally, language use and phonetic aptitude did not influence L2 acquisition or L1 attrition of vowels shared by both Arabic and English.

Chapter Four (Study III)

L2 acquisition and L1 attrition of VOT of word-initial voiceless plosives in highly-proficient late bilinguals: Exploring the role of phonetic aptitude and language use

Abstract

While much research has examined second-language (L2) acquisition, little research has examined language (L1) attrition in terms of the VOT of voiceless stops, and few studies have investigated the effect of individual differences in late bilinguals who learned a language later in life, in terms of the L2 acquisition and L1 attrition of segments. The current chapter examines L2 acquisition and L1 attrition in the VOT of word-initial voiceless stops among late E-A and A-E bilinguals. This chapter also investigates whether late L2 learners' phonetic aptitude and language use affect target-like learners in their VOTs of voiceless plosives of both English and Arabic. The data were obtained from 60 participants divided into four groups: 15 monolingual Arabic speakers, 15 monolingual English speakers, 15 E-A bilinguals and 15 A-E bilinguals. The bilinguals had been living in their L2 environment for an average of 20 years. The participants narrated five different cartoons: two in Arabic and three in English, and the bilinguals narrated them in both Arabic and English. In addition, both groups of bilinguals completed a proficiency test, a phonetic aptitude (talent) test and a language background questionnaire. VOT was measured using Praat from the interval between the plosive release and the onset of voicing. The monolingual groups' results revealed that the VOT of voiceless plosives differed in Arabic and English. Additionally, I found that the E-A bilinguals' English /k/ differs reliably from that of monolingual English speakers is broadly in line with the SLM's assumed flexibility of phonetic categories in the L1. Moreover, the results for A-E bilinguals' acquisition of /p/, but not /t/ or /k/, in the present chapter support the SLM theory. Finally, I found no evidence that language use or phonetic aptitude influenced the A-E and E-A bilinguals' productions in terms of L1 attrition and L2 acquisition.

Keywords

Bilinguals, second-language (L2) acquisition, first-language (L1) attrition, voice onset time (VOT), voiceless stops, language use, phonetic aptitude (talent)

4.1 Introduction

In general, adults speaking a second language (L2) adopt properties of their first language (L1) in their L2 speech, and are therefore likely to be identified as non-native speakers (Scovel, 1969; Brennan et al., 1975; Flege, 1980, 1981). Living in an L2 country may affect the L1, causing the L2 to play a dominant role in everyday life and decreasing the use of the L1 and contact with other native speakers. This increase in L2 use can thus result in a reduction in L1 use, causing a change in the speaker's linguistic abilities in their L1. This phenomenon is known as L1 attrition (Freed, 1982; Schmid, 2007). When the domains of phonology or phonetics are affected by L1 attrition, L1 speakers may be identified as having a foreign accent. The current chapter combines L2 acquisition and L1 attrition research and presents production data from two groups of late bilinguals (A-E and E-A), who are highly fluent speakers of both languages and who differ in their phonetic aptitude and language use, to consider possible bidirectional L1–L2 influences in their phonetic systems.

The chapter draws mainly on the SLM (Flege, 1995) to explain a bidirectional L1–L2 influence on a bilingual's speech. The SLM posits a process known as 'equivalence classification', which means that bilinguals classify L2 sounds that are similar to one of their L1 sounds as the L1 sound, thus essentially failing to differentiate between the L1 sound and the L2 sound. As a result, the L2 sound is produced in a non-native manner because it is produced like the L1 sound. Providing that L2 sounds differ sufficiently from L1 sounds, new categories can be established in the L2. However, within a bilingual's L1–L2 phonetic space, new L2 categories may still diverge from those of monolingual native speakers, for example, to retain contrasts with the L1 categories of the bilingual. Therefore, the production of the L2 speakers (who acquired new L2 categories) might still diverge from the speech of a native speaker. The SLM hypothesises that bilinguals have a common space with regard to their L1 and L2 phonetic categories and that this remains flexible in adulthood to some degree. The SLM also assumes that phonetic systems remain flexible over the lifespan, implying that L1 categories can change due to the influence of L2 acquisition, leading to a foreign accent in the L1. For this reason, the SLM has been adopted to interpret phonetic L1 attrition presented in previous research (de Leeuw, 2009; Chang, 2012; Mayr et al., 2012; Bergmann et al., 2016). Understanding how phonetic categories are organised in a speaker who uses two languages requires characterising phonetic properties in both the L2 and L1 speech (Flege & Eefting, 1987a, 1987b; Mennen, 2004; Chang, 2012; de Leeuw et al., 2012, 2013; Mayr et al., 2012). This chapter focuses on L2 acquisition and L1 attrition in the production of VOTs, particularly the voiceless plosives /p/, /t/ and /k/, in two groups of late bilingual A-E and E-A adults. VOT is the time between the burst of a stop occlusion and the onset of voicing, which characterises the contrast of voicing in most languages for initial stop consonants (Lisker & Abramson, 1964).

To examine the timing of voicing in plosives, the VOT has been applied in studies of many languages and has become one of the most important methods to distinguish between voiced, voiceless, aspirated and unaspirated plosives (especially in word-initial position). Voiceless stops are the focus of this study because the acoustic cue of VOT contributes to the perception of a foreign accent in L2 speakers and L1 attriters (Flege, 1984; Flege & Eefting, 1987b; Major, 1987; Sancier & Fowler, 1997; Schoonmaker-Gates, 2015; Tobin et al., 2017). This study measures the production of VOT by monolingual native speakers of both languages and examines the relationship between L1 attrition and L2 acquisition in highly fluent, late consecutive A-E and E-A bilinguals.

4.2 Literature Review

4.2.1 VOT Definition

As mentioned before in section 4.1, Lisker and Abramson (1964, p.422) defined VOT as “the time interval between the burst that marks the release of the stop closure and the onset of quasi-periodicity that reflects laryngeal vibration”. There are three categories of VOT: (1) voicing lead or prevoicing, which comprises negative VOT values, where the vocal folds start vibrating before the burst of the stop; (2) short-lag VOT with small positive VOT values or values around zero; and (3) long-lag VOT or aspirated with large positive VOT values. Figure 4.1 (adapted from Deuchar & Clark, 1996, p.25) displays the Arabic VOT range; it shows that the phonemically voiceless stops in Arabic almost fall into the phonemically voiced range of English stops, while the Arabic stops that are phonemically voiced and the English phonemically voiceless stops are at opposite ends of the continuum (see Figure 4.1). In other words, the voiced plosives /b/, /d/ and /g/ are produced with a short lag in English, but with lead voicing in Arabic, and the voiceless plosives /t/ and /k/ are produced with a long lag in English and a short lag in Arabic (Arabic has no /p/). As illustrated in Figure 4.1, if the duration between the burst of the stop and the onset of the following vowel is 30 ms or less, it is considered 'short-lag', and if the duration is more than 30 ms, it is considered 'long-lag', or aspirated (Deuchar & Clark, 1996).

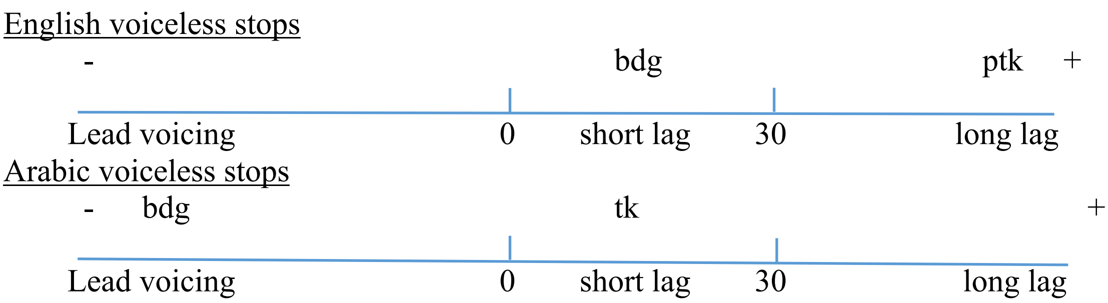


Figure 4.1. Voiced and voiceless stops in English and Arabic (adapted from Deuchar & Clark, 1996, p.25)

Although VOT is a frequently used measure in linguistic studies, its reliability to distinguish between voiced and voiceless stops has been questioned. For example, Cho et al. (2019) have recently argued for a need of multi-dimensional approaches to understand voicing contrasts across languages. Nevertheless, this study uses VOT as a measure because its frequency in the literature allows for comparisons with previous studies.

Bohn and Flege's (1993) findings suggested that VOT may not be as critical to the perception of stop voicing as was supposed. Docherty (1992) also claimed that VOT neglected stops in word-final and word-medial positions, and Garamazza et al. (1973) claimed that VOT is an 'insufficient' cue to the voicing contrast for French-English bilinguals. Important acoustic cues in English other than VOT include low frequency energy in following vowels, burst loudness, fundamental frequency, pre-voicing and segmental duration (Klatt, 1975, p.695). Despite these limitations, VOT remains a central acoustic parameter in linguistic studies measuring word-initial stops.

4.2.2 Acquisition of L2 Voiceless Stops in Arabic and English

Influences of the L1 on the L2 have been attested in bilinguals who speak two different languages that have contrasting VOTs. While some research has considered the acquisition of VOTs in Arabic and English (Flege, 1987, 1991; Flege & Eefting, 1987a, 1987b; Simon, 2009; Simon & Leuschner, 2010; Schmid et al., 2014; Stoehr et al., 2017; Ahn, 2019; Kim et al., 2018; Cho et al., 2019), few studies have examined VOTs in L2 English productions by Arabic adult speakers (Flege, 1980, Flege & Port, 1981; Port & Mitleb, 1983). In addition, to the best of my knowledge, no previous study has examined VOTs in L2-Arabic productions by E-A bilinguals. Voiceless plosives /p/, /t/ and /k/ are acquired in three different patterns by bilinguals who learned L2 aspiration in languages such as English and German and whose first language is a pre-voicing language such as Arabic, Spanish or Dutch.

Stoehr et al. (2017, p.486) summarised the three different acquisition patterns. The first is native-like acquisition, which was observed in highly advanced learners and phonetically trained participants (Simon & Leuschner, 2010), and in which the participants produce the VOT of plosives in the L2 the same as a native speaker. The second pattern is differential acquisition, which was observed in phonetically untrained participants (Simon & Leuschner, 2010) and in those with some level of L2 proficiency, and where participants produced the VOT of the L2 plosives with some differences to native speakers. The last pattern is a complete transfer from L1-to-L2, where the participant produces the L2 plosives as if they were speaking their L1 (Flege & Port, 1981; Flege, 1987; Flege & Eefting, 1987a, 1987b; Flege, 1991; Simon, 2009; Simon & Leuschner, 2010; Schmid et al., 2014).

Several studies have found instances of near-native acquisition in bilinguals in terms of VOT (Simon, 2009; Simon & Leuschner, 2010; Schmid et al., 2014). Simon (2009) and Simon and Leuschner (2010) examined the acquisition of VOT in highly proficient L2 learners of English who are native speakers of Flemish. Simon (2009) examined native Flemish speakers' L2 acquisition of the English laryngeal system, mainly aiming to test whether the Flemish (a voicing language) L1 native speakers would successfully acquire and produce the short-lag stops of the L2 (English, an aspirating language) since short-lag plosives reportedly occur early in L1 acquisition. Additionally, short-lag stops "can be considered unmarked as one member of the contrast is formed by short-lag stops in both voicing and aspirating languages" (Simon, 2009, p.377). Furthermore, Simon's (2009) study aimed at testing whether L1 native speakers would acquire the long-lag stops of the L2, since aspiration is a salient realisation acoustically. By examining natural conversations and controlled reading tasks, Simon (2009) found that native speakers of Dutch did not acquire the English short-lag stops, but they did acquire the English long-lag aspirated stops. As with Dutch, the Arabic language (MSA) contrasts short-lag with pre-voiced stops, but English has a contrast between long-lag and short-lag stops.

Similarly, Schmid et al. (2014) examined 20 advanced Dutch-English bilinguals and nine native English speakers and found native-like VOT acquisition patterns. All the bilingual participants were English students or English teachers at Dutch-language higher education institutions and at the same level of proficiency. The authors explored VOT and vowel discrimination, and they rated the participants according to their global native-like speech. The results revealed that no difference existed between the two groups of participants with regard to VOT and that both groups had native-like VOTs for English plosives.

Other studies have found differential acquisition in terms of VOT, such as Flege (1987, 1991), Flege and Eefting (1987a, 1987b) and Simon and Leuschner (2010). The differences in VOT acquisition patterns occur when bilinguals produce VOT differently in their L2 than they do in their L1, but still deviate from native speakers' VOT in the L2. These differences have been observed in L1-Spanish bilinguals who learned L2 English. Flege (1991) found that the bilinguals' VOT productions were longer in English than they were in Spanish but that their English VOTs were nevertheless shorter than those of monolingual English speakers. A similar pattern emerged in bilinguals with L1-Spanish who learned English as an L2 during childhood, regardless of whether they were immersed in an English environment or not (Flege & Eefting, 1987a). Comparable results were found in the Netherlands using Dutch native speakers who received informal instruction in their L2-English and L3-German phonetics (Simon & Leuschner, 2010). Distinct VOT values in the speakers' productions in Dutch were reported to have short-lag voiceless plosives compared to the aspirated voiceless plosives in English and German; however, they still appeared shorter than the VOTs of English and German monolinguals even though no direct statistical comparison was administered between the German (their L3) and English (their L2) and the monolinguals. More specifically, both trained and untrained participants acquired the aspiration, but they transferred the Dutch pre-voicing to German (their L3) and English (their L2). Level of L2 proficiency can play a role in L2 productions and bilinguals can distinguish L1 and L2 stops in terms of VOT but still not attain native-like VOT values in their L2.

A few studies have found instances of complete VOT transfer from L1 to L2 in bilinguals (Flege & Port, 1981; Flege, 1987). For example, Flege and Port (1981) observed a complete VOT transfer from L1 to L2 in native Arabic learners of English as an L2, who had lived in the US for about two years. Flege and Port (1981) examined six adult male Arabic native speakers' VOTs in the stop voicing contrast (voiceless vs voiced stops) in both word-final and word-initial position in CVC minimal pairs in both Najdi Saudi Arabic and English. The results revealed that the participants' English VOTs of /t/-/d/ and /k/-/g/ were similar to their Arabic ones and were shorter than the VOTs of English monolinguals. Native Arabic learners of English showed no signs of phonetic differences in the VOT between L1 and L2. However, the /p/ timing suggested that the Saudi participants grasped the phonological nature of the English /p/, which is probably because there is no /p/ in Arabic and thus no L1 influence on the L2 /p/.

To summarise, previous studies of VOT addressed the acquisition of voiceless plosives. Native-like acquisition of voiceless short-lag stops has only been found in a few studies, such as Flege (1987) and Stoehr et al. (2017). Moreover, L2 acquisition and L1 attrition in the production of voiceless plosives in Arabic and English by highly proficient late consecutive A-E and E-A bilinguals has not been reported. The current research therefore fills a gap in the literature by analysing data on the production of voiceless plosives in Arabic and English by native speakers of Arabic and English as well as bilinguals.

4.2.3 L1 Attrition of Voiceless Stops in Arabic and English

Influence of the L2 on the L1 can be observed in L2 speakers in the domain of phonetics (Flege, 1987; Flege & Eefting, 1987; Major, 1992; Mayr et al., 2012; Stoehr et al., 2017). Bilinguals whose dominant language is the L2 are more likely to show L1 attrition than are those whose dominant language is the L1 (Schmid & Köpke, 2007). The current study also investigates the VOTs of Arabic and English voiceless plosives in A-E and E-A bilinguals, who might show attrition in their L1.

Limited research has been done on the L1 attrition of VOT. Phonetic attrition can be a drift in the VOT values of the L1 towards the VOT values of the L2. A few studies have examined attrition in terms of VOT in the productions of highly proficient L2 speakers (Flege, 1987; Major, 1992; Sancier & Fowler, 1997; Mayr et al., 2012). Major (1992) examined the VOT of the sounds /p/, /t/ and /k/ in the speech of late English-Portuguese bilinguals and found that all the participants showed attrition in their native English to differing degrees (Major, 1992). For example, the participants' productions of voiceless stops in English and Portuguese differed from the native speakers' VOT productions. The productions of English stops moved in the direction of the VOTs of Portuguese native speakers, and vice versa. Moreover, the results showed evidence that proficiency in the L2 was correlated with loss in the L1. VOT measurements showed that the highly proficient L1 English L2 Portuguese speakers had a higher degree of deviation from the productions of

English native speakers than did the low L2 proficiency speakers. Flege (1987) identified a similar bidirectional influence of the consonant /t/ among American and French speakers, who had lived in the L2 environment (in France or America, respectively) for many years. In addition, Flege and Eefting (1987) conducted a study of the /t/ phoneme in Dutch-English bilinguals and found that bilingual L1 Dutch speakers produced the Dutch /t/ with VOT values that were shorter than the productions of monolinguals, making the VOT even shorter than the Dutch norm, which is already shorter than the English norm. This modification was particularly prevalent in speakers who were highly proficient in English, and it represents a shift not towards the English norm, but away from it. This serves to differentiate explicitly between the Dutch and the English /t/ phoneme.

Mayr et al. (2012) also found evidence for L1 attrition of VOT. As mentioned in Chapter Three, Mayr et al. (2012) investigated the production of plosives and vowels in two late consecutive Dutch-English bilinguals, where one bilingual had moved to an English environment. They compared the level of L1 attrition in 62-year-old monozygotic twin sisters who were bilingual in Dutch and English, one of whom (MZ) had moved to the UK at the age of 32. Both twins used Dutch and English daily, but MZ used English more often than she used Dutch, while her counterpart in the Netherlands (TZ) did the opposite. The authors observed systematic differences between the two speakers that affected some phonemes and not others, suggesting that attrition may not affect all areas of pronunciation equally. For example, while MZ produced the Dutch voiced plosives /b/ and /d/ as would a native speaker, the Dutch voiceless plosives produced by MZ had longer VOT values than those produced by TZ. MZ's VOT values fell between the norms of the native Dutch and the native English. Overall, Mayr et al. (2012) found attrition in the voiceless plosives, but not the voiced plosives, suggesting that some areas of pronunciation might be more sensitive to attrition than others.

A recent study that found L1 attrition of VOT was conducted by Stoehr et al. (2017), and to the best of my knowledge, it is the only study that has examined L2 attainment and L1 attrition with regard to the VOT system. The authors examined Dutch-German (N = 18) and German-Dutch (N = 23) late bilinguals who lived in the Netherlands, and two control groups comprising 27 German monolinguals and 29 Dutch monolinguals. Their results indicated that complete L2 engagement might be useful in acquiring the L2; however, it might cause phonetic attrition in the L1. More specifically, German-Dutch bilinguals acquired the Dutch

(L2) VOT short-lag for /p/ and /t/ in a native-like way but did not reach the native speakers' proficiency in terms of pre-voicing, such that German-Dutch bilinguals did not produce the Dutch (L2) /b/ and /d/ in a native-like way. The VOT of their native German seemed to be partly affected "by language attrition as revealed by shorter than monolingual-like VOT in voiceless plosives" (Stoehr et al. 2017, p.503). By contrast, language attrition did not affect the voiced plosives. The results of the Dutch-German bilinguals revealed that the participants produced longer VOTs for German voiceless plosives than they did in their native Dutch, while they pre-voiced the voiced plosives in both German (L2) and Dutch (L1). The results also illustrated that phonetic space of their L1 and L2 shows absolute phonological difference between voiced and voiceless stops, while the differences in the German and Dutch languages are "present for voiceless plosives, but absent for voiced plosives" (Stoehr et al., 2017, p.502).

In summary, a number of studies have proposed a bidirectional influence of the L1 and L2 segmental sound systems that is measurable at a fine-grained level of detail when languages differ in their voiceless and voiced plosives.

4.2.4 Factors that Influence L1 Attrition and L2 Acquisition

Chapters Two and Three explored factors which influence L2 acquisition and the degree of L1 attrition in the phonetic domain – in particular, language use and phonetic aptitude. These factors influenced L2 acquisition and the degree of L1 attrition in the prosody of *wh*-words in A-E and E-A bilinguals (see Chapter Two for more detail) but did not show the same influences in L2 acquisition and the degree of L1 attrition in the formants of shared vowels among the same bilinguals (see Chapter Three for more details). Hence, these factors may or may not affect or play a role in the production of A-E and E-A bilinguals' VOT of the voiceless stops.

The Role of language use in L2 Acquisition and L1 Attrition

Previous studies have examined how the frequency of L2 use, particularly with native speakers, affects language acquisition (Purcell & Suter, 1980; Leather, 1987; Shen, 1990; Thompson, 1991; Flege & Fletcher, 1992; Flege, 1995; Moyer, 1999; Guion et al., 2000; de Leeuw, 2009; Mayr et al., 2012; Stoehr et al., 2017). A study by Stoehr et al. (2017) indicated that complete L2 immersion may be advantageous in L2 acquisition but that it may also cause

L1 phonetic attrition. Whereas previous studies have shown that language use has a small influence on L2 acquisition, other studies have revealed no obvious impact from L2 use or input (see Chapter Two, Section 2.2.3.1 for more information). Generally, it appears that certain kinds of language use – for example, in the home or in situations where code-switching is unlikely – have an influence on L1 attrition. Notably, as most of these studies concern a global foreign accent, it is unclear which aspects of pronunciation may be affected by language use. To the best of my knowledge, no previous study has directly considered how language use affects L2 acquisition and L1 attrition of VOT in plosives, except Stoehr et al. (2017), who studied L1 attrition and L2 attainment with regard to the VOT system. Their results suggested that complete L2 engagement might be beneficial in obtaining the L2; however, it might cause phonetic L1 attrition. That said, Sancier and Fowler (1997) provided some indirect evidence that language use may affect L2 acquisition and L1 attrition. Specifically, they found evidence that their bilingual L1 Brazilian Portuguese L2 English speaker's VOT values in both the L1 and L2 shifted in the direction of the ambient language as the speaker spent time either in the L1 or L2 environment over the course of the study (Sancier & Fowler, 1997). Assuming that the speaker would more frequently use the L2 in the L2 environment and the L1 in the L1 environment, these VOT shifts may be related to language use (see also Tobin et al., 2017). In addition, Mayr et al.'s (2012) study of the monozygotic twins suggested that long-term L2 experience affected L1 VOT production of voiceless plosives. To contribute to a more detailed picture of the role of language use in L2 acquisition and L1 attrition in the phonetic domain, the current study focuses on one particular aspect of VOT values. Therefore, the present study's objective is to offer a more in-depth examination of the effect of language use on the VOT values of voiceless stops in L2 learning and L1 attrition.

The Role of Aptitude in L2 Acquisition and L1 Attrition

Very few studies have focused on the area of phonetic aptitude and examined its links to L2 acquisition and L1 attrition (Tahat et al., 1981; Flege et al., 1995; Thompson, 1995). To the best of my knowledge, no previous study has directly considered how phonetic aptitude affects L2 acquisition and L1 attrition of VOT in plosives. Using proficiency tests, Hopp and Schmid (2013) found an indirect link between aptitude and the presence of a foreign sounding accent. However, self-rating or indirect measures are less reliable than direct measurements of language aptitude (Jilka et al., 2010; Hinton, 2012).

Mennen's (2004) study attempted to make clear that some participants are simply 'outliers' from the general population, in that they have the ability to achieve native-like production of both their L1 and L2. This issue raises the following question: What makes these individuals particularly good at learning the L2 while maintaining nativeness in their L1? One possibility is that participants who do well in the L2 – and who also show no effects on the L1 – may be phonetically talented, which could explain their success. The present chapter uses specific aptitude tests to investigate the effect of phonetic aptitude on the production of voiceless stops in L1 attrition and L2 acquisition.

4.2.5 Arabic and English VOT Systems

As previously mentioned (see sections 4.2.2 and 4.2.3), many studies have found the VOT values of voiced and voiceless stops to be important acoustic cues in a number of languages. Lisker and Abramson (1964) analysed the VOT of 11 languages, including English. Each stop category falls into one of the three ranges mentioned in section 4.2.1: negative VOT (pre-voicing), short-lag VOT and long-lag VOT. Some languages have plosives with short-lag VOT and long-lag VOT, others have stops with short-lag VOT and pre-voicing negative VOT, and still others make a three-way distinction with pre-voicing, short-lag and long-lag VOT. Moreover, Cho. et al. (2019) mention that some languages have four- and five-way stop distinctions, for example Jangli and Urdu, which have a four-way contrast, and Sindhi and Siraiki, which have a five-way contrast. English (British English) has only short- and long-lag stops (Lisker & Abramson, 1964), whereas Arabic has the short-lag plosives /t/ and /k/ (which are long-lag plosives in some Arabic dialects, but not as long as those in English; Alanazi, 2018) and the pre-voiced stops /b/, /d/ and /g/. Many researchers (Lisker & Abramson, 1964; Klatt, 1975; Port & Rotunno, 1979; Weismer, 1979; Keating et al., 1983; Docherty, 1992) have examined English VOT values.

There is little research on Arabic VOT for word-initial plosives (Flege, 1980; Khattab, 2000; Alanazi, 2018). Arabic and English differ in their VOT patterns, but both have two stop categories (Lisker et al., 1964). English is often described as having three voiceless stops—/p/, /t/ and /k /—and the three voiced stops /b/, /d/ and /g/. According to Lisker and Abramson (1964), the stops /b/, /d/ and /g/ are voiced in medial positions; however, in initial positions both voiceless and voiced stops appear to be actually voiceless as both are produced with a silent closure. Thus, based on Lisker and Abramson (1964), it is important to examine the aspiration of stops in initial positions to distinguish between /p/, /t and /k/ and /b/, /d/ and /g/. In that sense, the voiceless stops /p/, /t/ and /k/ are characterised as having long-lag VOT patterns, while the plain/voiced stops have short-lag VOT patterns (Lisker & Abramson, 1964).

Although little research has dealt with the issue of VOT in different Arabic dialects (Al-Ani, 1970; Yeni-Komshian, Garamazza, & Preston, 1977; Flege & Port, 1981; Port & Mitleb, 1989; Alghamdi, 1990; Radwan, 1996; Jesry, 1996; Mitleb, 2009; AlDahri, 2013), studies that have been carried out mostly agree that voiced stops in Arabic are pre-voiced (except Mitleb, 2009, and AlDahri, 2013) and that voiceless stops fall into the short-lag range of the continuum. Notably, not all of the studies reviewed below reported which specific dialect they studied, but it seems that they used MSA rather than colloquial or national dialects, except for Khattab (2002), Flege and Port (1981) and Alghamdi (1990), who used colloquial dialects in their studies.

One of the earliest studies in Arabic phonology was conducted by Al-Ani (1970), who measured the duration of aspiration of voiceless stops by recording himself reading lists of words. He found that the VOT productions of /k/ were between 60 and 80 msec, while the VOT values of /t/ were between 30 and 40 msec (Al-Ani, 1970). These values were at the low end of the long-lag range rather than in the short-lag range.

Yeni-Komshian et al. (1977) studied the production of stops in MSA by asking eight Lebanese adults to read words and sentences with stops in the context of the three short vowels /a/, /iː/ and /ʊ/. The researchers found that voiced stops were pre-voiced in all participant productions, while the voiceless ones fell within the short-lag region (Yeni-Komshian et al., 1977). They also found no major differences between stops produced in words in isolation and words produced in sentences (Yeni-Komshian et al., 1977). Moreover, they found that there was a tendency for shorter negative VOT and longer short-lag VOT in the production of stops before /iː/ than before the other vowels /a/ and /ʊ/ (Yeni-Komshian et al., 1977).

Jesry (1996) compared the Arabic voicing contrast in MSA. The study comprised three adult Syrians who read target words containing word-initial stops and fricatives followed by one of the following vowels: /ɪ/, /iː/, /a/, /ɑː/, /ʊ/ and /uː/. Jesry's (1996) results showed that the voiceless stops fell within the short-lag region, while the voiced stops were pre-voiced in all of their participants' productions, similar to the results obtained by Yeni-Komshian et al. (1977).

Several researchers have investigated English VOT patterns (e.g. Lisker & Abramson, 1967; Klatt, 1975; Docherty, 1992; Khattab, 2002; Scobbie, 2002). Docherty (1992) studied the VOT patterns of SSBE and reported in detail the various features of the timing of voicing in voiced and voiceless obstruents. Docherty's (1992) participants were five adult male British speakers of SSBE; they were all educated and brought up in South East England. He measured the VOT of stops and fricatives in different contexts (in isolation and in a carrier phrase). Voiceless stops in initial positions (not following /s/) were aspirated in British English as they were in the long-lag range: 46 msec for /p/, 66 msec for /t/ and 66 msec for /k/ in words in isolation. Voiced stops exhibited shorter VOT values (25 msec) for /b/, but with slight aspiration for /d/ (33 msec) and for /g/ (40 msec). Voiceless unaspirated stops were also in the short-lag range, but with shorter VOT values than the voiced stops in words in isolation. Docherty (1992) stated that even pre-voicing was recorded for some voiced stop tokens, particularly of /b/ and /d/.

Table 4.1 shows the average VOT values of the stops /t/ and /k/ for Arabic, as reported by Al Dahri (2010), Alghamdi (2006) and Mitleb (2001). Table 4.2 displays the average VOT values for MSA from Yeni-Komshian et al. (1977) and Jesry (1996). Table 4.3 shows the VOT values of the voiceless stops for English from Lisker and Abramson (1964; American English) and Docherty (1992; British English). Tables 4.1, 4.2 and 4.3 show that the voiceless stops in both languages have positive VOTs, with English having larger positive VOT values than does Arabic.

Table 4.1. Average VOT Values for /t/ and /k/ for Arabic in Different Dialects

Plosive	**Al Dahri and Alotaibi, 2010 (Arabic, Najdi dialect)**	**Alghamdi, 2006 (Arabic, Gamdi dialect)**	**AlDahri and Alotaibi, 2010 (Quranic Classical Arabic)**
	Mean	**Mean**	**Mean**
/t/	49	39	36
/k/	52	42	37

Table 4.1. Average VOT Values for /t/ and /k/ for MSA

Plosive	Yeni-Komshian et al., 1977 (MSA)			Jesry, 1996 (MSA)
	Mean			Mean
	/a/	/ʊ/	/i/	
/t/	20	25	30	27.82
/k/	25	30	30	32.19

Table 4.2. Average VOT Values for /t/, /k/ and /p/ for English in American and British Dialects

Plosive	Lisker & Abramson, 1964 (AE)		Docherty, 1992 (BE)	
	Mean	Range	Mean	Range
/t/	70	30-105	64	30-110
/k/	80	50-135	62	30-150
/p/	58	20-120	42	10-80

To align with the study's purpose, I focus on SSBE (Doherty, 1992) and MSA (Yeni-Komshian et al., 1977; Jesry, 1996), and I exclude AlDahri and Alotaibi's (2010) and Alghamdi's (2006) data because they used the colloquial Saudi dialects Najdi and Gamdi. The average VOT values given in Table 4.2 for the MSA /t/ range from 20 to 30 ms, while the average VOT value for British English /t/ is 64 ms. The average VOT value of the Arabic /k/ ranges from 25 to 32 ms, whereas the average VOT value for the English /k/ is 62 ms. As I mentioned previously, there is no /p/ sound in Arabic, so there is no average value for it. The VOT value for the English /p/ is 42 ms.

Because AlDahri and Alotaib's (2010) and Alghamdi's (2006) studies used the Saudi dialects Najdi and Gamdi rather than MSA for Saudi Arabic /k/ and /t/, the VOT values did not fall exactly into the short-lag region but were in the low end of the long-lag region, bordering on short lag. The /t/ and /k/ sounds in these Saudi dialects are thus considered to be aspirated (Flege & Port, 1981).

Table 4.4 presents the generalisations about Arabic and English VOT patterns provided by Khattab (2002). Note that the results from some previous studies are inconsistent with these generalisations, notably AlDahri and Alotaibi (2010), Mitleb (2001) and Scobbie (2002).

Table 4.3. Generalisations about Arabic and English VOT Patterns in Word-Initial Positions Provided by Khattab (2002, p.218)

	Arabic VOT patterns	**English VOT patterns**
Initial voiced stops	predominantly lead voicing (VOT between -60 and -90 msec).	unaspirated (VOT between 0 and 25 msec) or voiced
Initial voiceless stops	delay of between 25 and 60 msec in voicing, relative to the release of the stop	delay of between 50 and 80 msec in voicing, relative to the release of the stop
Presence or absence of vocal fold vibration in the closure duration	contrastive	not contrastive

Arabic and English have important differences in their VOT patterns. In English, before and during the production of the plosives /b/, /d/ and /g/, there are only rarely vocal cord vibrations, such that /b/, /d/ and /g/ are typically considered to be voiced and unaspirated (but cf. Scobbie, 2002, and a number of English regional dialects). Additionally, in English [p^h], [t^h] and [k^h] are strongly aspirated. Therefore, the contrast between voiceless and voiced plosives is in aspiration in English, but in the absence or presence of glottal pulsing in Arabic (cf. Flege & Port, 1981; Khattab, 2002).

The empirical evidence in some of the previous studies (Yeni-Komshian et al., 1977; Jesry, 1996; Docherty, 1992) suggests that the VOT in Arabic and English plosives differs sufficiently. The SLM therefore predicts that bilinguals would establish a new category for the L2 sounds and produce the L2 sounds differently from the L1 sounds, and they would thus potentially be native-like in their L2 productions. As Arabic (MSA) does not have a /p/ sound, I would expect that highly fluent A-E late bilinguals would form a new category and produce the English /p/ similarly or close to the production of /p/ by native English speakers. Furthermore, both Arabic and English have voiceless stops /t/ and /k/, even though English /p/, /t /and /k/ are produced in the long-lag range, while the /t/ and /k/ in Arabic are produced in the short-lag range. I would therefore expect that highly fluent A-E and E-A bilinguals would create a new category for the voiceless stops /t/ and /k/ in English and Arabic and produce them with VOT values similarly or close to the production of native speakers (Arabic and English). In addition, the SLM also assumes that phonetic systems remain flexible over the lifespan, implying that L1 categories can change due to the influence of L2 acquisition, leading to a foreign accent in the L1. Hence, I expect that E-A and A-E bilinguals would produce voiceless plosives /t/ and /k/ with compromised VOT values in both languages.

In addition, I explore how the amount of L2 language use and the phonetic aptitude affect both L1 attrition and L2 acquisition in their phonetic systems. The present investigation adds to the existing VOT literature on L2 acquisition and L1 attrition and combines investigations of L1 attrition and L2 acquisition in the same speakers.

4.3 The Current Study

The main aim of the present study is to investigate whether highly fluent late consecutive A-E bilinguals who moved to the UK in adulthood and highly fluent late consecutive E-A bilinguals who moved to the KSA or Yemen in adulthood show native-like voiceless plosive productions in their L2s and/or show L1 attrition for voiceless plosives. This chapter investigates the VOT of voiceless stops as the phonetic variable in the L1 and L2 speech of A-E and E-A bilinguals. The VOT of voiceless stops was selected because the literature indicates differences between English and Arabic in the VOT of voiceless plosives. Moreover, few studies have been conducted on the production of voiceless plosives in Arabic. The current study attempts to fill this gap in the literature. The present study also attempts to test whether conclusions drawn from previous studies hold true across two typologically different language groups, such as Arabic and English.

The second aim of the present study is to determine whether language use has an effect on L1 attrition and L2 acquisition of the VOT of the voiceless plosive. Many researchers have assumed that exposure to the L2 and prolonged use leads to L1 attrition; hence, language use was selected as one of the study's main factors. All participants of this study were late L2 learners who had lived in the L2 environment for a substantial amount of time and thus had extensive exposure to the L2. Language use was measured using a detailed questionnaire.

Finally, the last aim of the current study is to ascertain whether phonetic aptitude plays a role in L1 attrition and L2 language acquisition in terms of the VOT of voiceless stops, using methods by Jilka (2007, 2009) and Jilka et al. (2008, 2011). In particular, aptitude was measured via a variety of phonetic aptitude tests. More native-like pronunciation in both the L1 and L2 may be found in participants with high phonetic aptitude, as suggested in Chapter Two.

4.5 Hypotheses

This study assesses L1 attrition and L2 acquisition in the production of plosives in late E-A and A-E bilinguals (i.e. bilinguals who speak two typologically different languages). It also examines whether late L2 learners' phonetic aptitude and language use influenced the production of plosives in Arabic and English. The current study explores the following hypotheses:

1. Arabic and English VOT: Based on the existing literature, I would expect a difference in the pronunciation in terms of VOT for voiceless plosives in Arabic (MSA) and English:

1a. The VOT of the voiceless stop of the Arabic monolingual speakers will be in the short-lag range.

1b. The VOT of the voiceless stop of the English monolingual speakers will be in the long-lag range.

2. L2 acquisition:

2a. As Arabic and English voiceless stops differ, I would predict that, based on the SLM, highly fluent learners of English or Arabic will form new categories for /t/ and /k/ in English and Arabic and they will produce them with VOT values much the same as the native speakers of Arabic or English.

2b. The SLM would predict that A-E bilinguals would establish a new category for the L2 sound /p/. Hence, I would expect that A-E bilinguals' /p/ is similar to that of native English speakers.

3. L1 attrition: Based on the previous literature and the SLM theory, L1 production may be affected by the L2; thus, the participants may produce plosives with compromise VOT values for the voiceless stops /p/, /t/ and /k/ in both languages, and both groups may differ from predominantly monolingual native speakers. L1 attrition may result in values that are between the norms for the L1 and the L2. Compared to the native speakers of Arabic and English:

3a. A-E bilinguals will produce the VOTs for the voiceless Arabic plosives /t/ and /k/ in a manner similar to English, and these will therefore be longer than the VOTs of Arabic monolinguals; and

3b. E-A bilinguals will produce the VOTs for the voiceless English plosives /t/ and /k/ in a manner similar to Arabic, and these will therefore be shorter than the VOTs of English monolinguals.

4. Language use: Based on Sancier and Fowler (1997), I will tentatively assume that language use will influence L1 attrition and L2 acquisition; that is, bilinguals with more L2 use will show a more native-like pattern in their L2 compared to those with less L2 use.

5. Phonetic aptitude: While there is no consensus in the literature regarding the role of phonetic aptitude in L2 acquisition and L1 attrition, I tentatively assume that participants with high phonetic aptitude will show more native-like production of the VOT in both their L1 and L2.

4.6 Methods

4.6.1 Participants

The study investigated the VOT of voiceless stops in the same participants as in Chapter Three. For convenience, Table 4.5 repeats the participant characteristics summarised in Chapter Three.

Table 4.4. Summary of Participant Characteristics

Participant group	N (gender)	Mean age (SD)	Mean AOA (SD)	Mean LOR (SD)
Monolingual Saudi, native Arabic	15 (4 males, 11 females)	36.06 (SD = 6.91)	NA	NA

speakers				
Monolingual English native speakers	15 (2 males, 13 females)	40.93 (SD = 9.65)	NA	NA
A-E bilingual	15 (one male, 14 females)	39.4 (SD = 3.75)	19 (SD = 2.8)	20 years (SD = 2.6)
E-A bilingual	15 (four males, 11 females	33.66 (SD = 3.9)	16.7 (SD = 1.12)	17 years (SD = 3.9)

4.6.2 Materials and Procedure

Quasi-spontaneous Speech for Consonants (Segmental Production)

I collected spontaneous speech production data for Arabic and English voiceless plosives from the participants. I used spontaneous speech because it "reflects overall abilities the best, allowing especially representative impressions of fluency, speaking rate, choice of words, choice of prosodic patterns and segmental realizations" (Jilka et al., 2008, p.228). Bilingual and monolingual participants narrated short cartoons, each chosen to elicit a particular plosive. Each cartoon task includes one or more pictures for the target word. To prepare for the quasi-spontaneous speech task, a native speaker of each language was asked to narrate the cartoon first, to determine the feasibility of a bilingual participant narrating that cartoon and eliciting the target words.

Each bilingual participant narrated three separate cartoons in English, featuring the voiceless plosives /k/, /p/ and /t/ and two cartoons in Arabic, which included /k/ and /t/, since the Arabic language does not have /p/. Table 4.6 shows the target words of the English and Arabic voiceless stops: 'cow', 'ti-ger' and 'pe-nguin' for the three English cartoons, and /kalb/ ('dog') and /tem-sɑːħ/ ('crocodile') for the Arabic cartoons. Previous research concerning the effects of a following vowel on plosives' VOT has been inconclusive. While Lisker and Abramson (1967) found no effect of a following vowel on VOT, Klatt (1975), Weismer (1979), and Port (1979) found that the VOT of plosives were longer when followed by tense high vowels than for all other vowels. Therefore, none of the plosives in the target words preceded tense high vowels. To keep the two initial sounds in the target words as similar as possible across the two languages, each plosive was followed by the same or a similar vowel in each language. For example, the plosive /k/ was followed by the vowel /a/ (either as a monophthong or the first part of the diphthong /aʊ/) in both Arabic and English.

Table 4.5. Arabic and English Plosives Target Words

Plosives	**English**		**Arabic**	
	Target word	**Description**	**Target word**	**Description**
/k/	‘cow’	A cartoon (Larson, 1992) of a cow eating grass, then ringing the doorbell. The cow is then seen eating the grass again.	/kalb/ ‘dog’	A cartoon (Larson, 1992) of two men talking about a dog (in Arabic, /kalb/). Then a group of dogs is seen having their own conversation.
/t/	‘ti-ger’	A cartoon (Bill Watterson, 1993) including pictures of a boy talking to a tiger.	/tem-saːħ/ ‘crocodile’	A cartoon (Larson, 1992) of people on a boat; a crocodile (in Arabic, /tem-saːħ/) watches them.
/p/	‘pen-guin’	A cartoon (Larson, 1995) of a group of penguins on an iceberg, standing around a polar bear wearing a penguin mask.	N/A	N/A

The bilingual participants were asked to narrate five different cartoons – three in English and two in Arabic. Three cartoons including the target words for the English plosives /t/, /k/ and /p/ were narrated in English by the bilingual participants on one day. On a second day, the bilingual participants narrated two cartoons in Arabic. As there is no /p/ sound in Arabic, there was no target word for /p/ in Arabic. Each monolingual participant only narrated the cartoons designed for their native language. Thus, their participation involved only one session. All participants were recorded with a handheld Sony tape recorder (SONY

ICD PX333Digital Voice Recorder) in a quiet environment. I used cartoons because they encourage participants to choose the vocabulary that is suggested in the pictures and they are likely to produce more authentic spoken utterances than reading aloud or answering direct questions from another speaker (Seliger & Shohamy, 1989). In the current chapter, the VOT values of the voiceless word-initial stops are measured using Praat (Boersma & Weenink, 2016) as the duration of the interval between the plosive release and the onset of voicing (see figure 4.1). The first token for each participant (60 tokens for each plosive) was selected because most of the participants replaced the target nouns with pronouns after the first mention. None of the tokens had to be excluded due to dysfluencies or other irregularities.

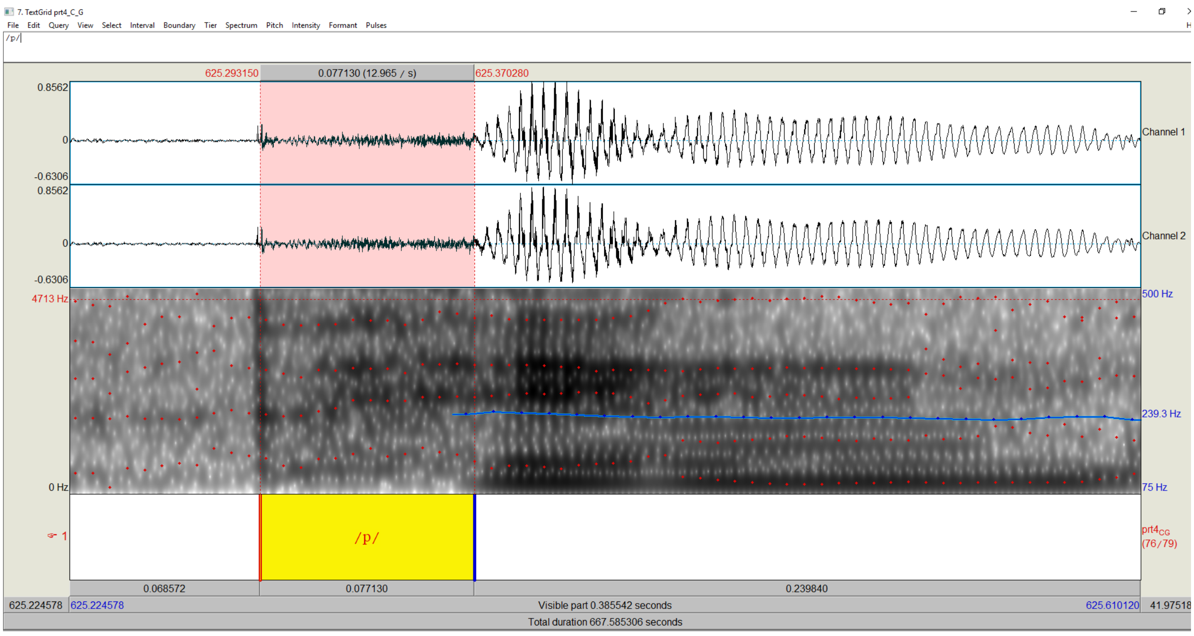


Figure 4.2. Example of a spectrogram and waveform of the /p/ by English native speakr

Background Questionnaire

The language background questionnaire described in Chapter Two aimed to collect sociolinguistic background information from the bilingual participants and provide answers to questions pertaining to the duration of contact with speakers of each language, the frequency and type of use of each language and the participants' attitudes towards the relevant languages and cultures**.** A more detailed description of the questionnaire is available in Chapter Two.

Language Proficiency and Aptitude

The results from the proficiency and aptitude tests were also adopted from Chapter Two. The TOEFL (computer-based version) and an APT were used to measure the bilinguals' Arabic and English language proficiency. See Chapter Two for more details.

Samples available on the Internet from MLAT Part III (Spelling Cues, which tests sound-symbol associations) and Part V (Paired Associates, which tests memory for novel words) and PLAB Part V (Sound Discrimination, which tests discrimination of pitch, orality and nasality) were adapted for the aptitude test. For a more detailed description of the aptitude test, see Chapter Two.

4.6.3 Overall Procedure

The snowball method was used for the data collection, which was conducted over six months in KSA and the UK, as mentioned in Chapter Two. After receiving the participants' consent via email, the bilinguals attended three sessions on separate days up to one week apart. The first session assessed participants' proficiency in the L2, followed by an aptitude test. The second and third sessions consisted of three different production tasks, with Arabic productions being recorded in one session and English productions in the other. The participants read lists of words and sentences six times, as described in Chapter Three. The participants then read the question-answer dialogs analysed in Chapter Two. Finally, voiceless plosive sound productions in Arabic (/k/ and /t/) and in English (/k/, /p/ and /t/) were recorded as the participants narrated short cartoons by Gary Larson (1992) using a handheld Sony tape recorder in a quiet environment. Monolingual participants only required two sessions, as they only participated in these production tasks in their native language.

4.6.4 Data Analysis

I conducted an analysis of the word-initial voiceless plosives in the participants' first productions of each target word. I first determined the beginnings and ends of the VOT part of the plosives in Praat and then labelled the voiceless plosives; both processes were done manually. A Praat script then extracted duration information (Boersma & Weenink, 2016). The VOT duration was measured from the interval between the plosive release and the onset of voicing. Next, a Praat script extracted VOTs for all the selected voiceless plosives. Mixed effects models were used to analyse the VOTs for each voiceless plosive. The mixed effects model allowed both fixed and random effects to be included in the analysis. I also used a linear model to see if aptitude or language use affected the VOT. I analysed the VOT values separately for each of the plosives (/k, /t/, and /p/). The analyses were performed using RStudio (Version 1.1. 456, 2018) and the lme4 package (Version 1.1-18-1, 2018). Depending on the analysis, the fixed effects in the analyses included language (Arabic vs English), group (A mono, E mono, A-E and E-A bilinguals), language use (high vs low) and aptitude (high vs low).

4.7 Results

4.7.1 Profile of Bilingual Participants

The A-E and E-A bilinguals who participated in this experiment were the same as the participants of the study described in Chapter Three. Table 4.7 (repeated here for convenience) summarises the proficiency, aptitude and language use among A-E and E-A bilinguals.

Table 4.6. Proficiency, Aptitude and Language Use among A-E and E-A Bilinguals

Measure	A-E bilinguals	E-A bilinguals	*t*-test *p*-value
Proficiency	TOEFL: mean = 42.1 (SD = 1.9; range = 39-45)	APT: mean = 43.7 (SD = 1.0; range = 42-46)	NA
Aptitude spelling cues	mean = 17.5 (SD = 4.2; range = 11-25)	mean = 19.8 (SD = 3.3; range = 14-24)	$t = -1.4494$ $p = .16$
Aptitude paired associates	mean = 14.6 (SD = 2.0; range = 12-18)	mean = 15.6 (SD = 2.2; range = 12-19)	$t = -1.4155$ $p = .17$
Aptitude sound discrimination	mean = 23.5 (SD = 2.3; range = 20-28)	mean = 23.5 (SD = 2.7; range = 20-28)	$t = 0.1883$ $p = .85$
Aptitude total	mean = 55.6 (SD = 7.8; range = 45-71)	mean = 58.9 (SD = 7.5; range = 49-69)	$t = -1.0437$ $p = .31$
Language use	More Arabic = 8 More English = 7	More Arabic = 5 More English = 10	NA

4.7.2 Comparison of Monolingual Speakers

This section presents the results of the comparison between English monolingual speakers (E mono), Arabic monolingual speakers (A mono) and A-E and E-A bilingual speakers in producing voiceless plosives (/k/, /t/ and /p/). The first analysis compares the patterns in the data for monolingual speakers for each voiceless plosive separately to determine whether the VOT values of the voiceless stops for monolingual Arabic speakers differ from those for monolingual English speakers. If they differ, the next step is to compare the data for bilingual and monolingual Arabic and English speakers to determine whether the bilingual groups differ from the monolingual ones. The final section examines whether significant differences that exist in the VOT values between the bilingual and monolingual groups for voiceless plosive production could be attributed to language use (more Arabic, more English) and phonetic aptitude (high and low).

VOT Values for the Voiceless Plosive /k/ in Arabic and English

Comparison of Patterns for Monolingual Speakers

First, the mean VOT for A mono is compared with that for E mono to determine whether the VOT of the voiceless /k/ differs between the two groups of monolingual participants. Figure 4.2 shows mean VOT values of /k/ in Arabic and English as produced by the A mono, E mono, A-E bilingual and E-A bilingual groups. The figure reveals that VOT values for E mono were numerically higher than those of A mono, meaning that E mono speakers tend to produce the /k/ in a more strongly aspirated manner (long lag) compared to the A mono group.

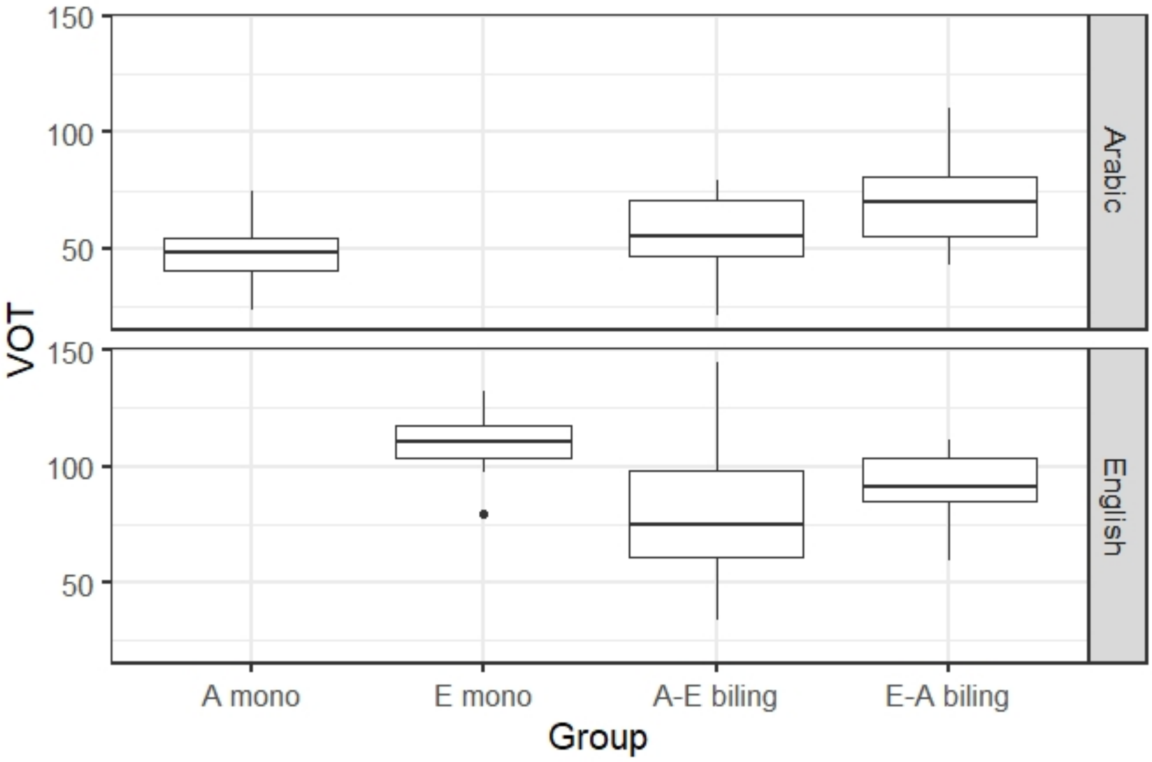


Figure 4.3. Boxplots for the VOT values (in msec) of the Arabic and English /k/ as produced by the A-E bilingual, E-A bilingual, A mono and E mono groups

To determine whether A mono and E mono speakers differ in a statistically significant way in terms of their VOT when producing the /k/ voiceless stop, I used mixed effects model. The response variable for this analysis is the VOT value in msec for each of the voiceless plosives /k/, /t/ and /p/. The fixed effect is group (A mono and E mono). Table 4.8 presents the results of this analysis, which reveal a significant effect of speaker group, showing that the VOT values of the /k/ phoneme were significantly higher (more strongly aspirated) for the E mono than for the A mono speakers.

Table 4.7. VOT Value Comparison between A Mono and E Mono Groups for the Voiceless /k/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	47.992	3.471	28	13.82	< .0001
Speaker group	163.163	4.91	28	12.9	< .0001

4.7.3 Comparison of Patterns for Bilingual Speakers

Comparison of VOT Values for E-A and A-E Bilingual Speakers and A Monolingual Speakers for /k/ in Arabic

Since Arabic and English monolingual participants differ in their production of the /k/ voiceless plosive, I then examined how bilingual speakers produce this phoneme in their Arabic productions and how they compare with Arabic monolingual speakers. Figure 4.2 above shows the numerical differences in the VOT values of the /k/ produced by the A mono, E-A bilingual and A-E bilingual groups. I used mixed effects model to determine whether the E-A and A-E bilingual groups shown in Figure 4.2 differ statistically significantly in terms of VOT. The results shown in Table 4.9 reveal that E-A bilinguals differ from A mono speakers in terms of the VOT values of the voiceless /k/, but A mono and A-E bilingual speakers show no significant differences. These results mean that the E-A bilingual group did not achieve target-like production of the voiceless stop /k/ in terms of VOT and that there is no evidence of L1 attrition in the A-E bilinguals.

Table 4.8. VOT Value Comparison of A Mono and E-A and A-E Bilingual Groups for /k/ in Arabic

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	47.992	4.43	42	10.84	< .0001
A mono vs A-E bilingual	8.49	6.263	42	1.36	= .182
A mono vs E-A bilingual	23.36	6.263	42	3.73	< .0001

To explain the significant differences between the E-A bilinguals and the Arabic monolingual speakers, I considered whether language use or phonetic aptitude influences E-A bilingual speakers' production of these phonemes. To determine whether aptitude and language use influence the VOT values in E-A bilinguals' Arabic productions of /k/, I used mixed effects models. Tables 4.10 and 4.11 summarise the statistical results and show that no significant difference exists between the aptitude groups and language use groups in terms of the VOT values of the voiceless /k/ in Arabic.

Table 4.9. VOT Value of /k/ in Arabic According to Aptitude Group

Fixed effects	**Estimate**	**Std. error**	***t*-value**	***p*-value**
Intercept	73.45	7.55	9.73	< .0001

Aptitude	-4.497	11.052	-0.41	= .691

Table 4.10. VOT Value of /k/ in Arabic According to Language Use Group

Fixed effects	Estimate	Std. error	*t*-value	*p*-value
Intercept	68.43	8.05	8.504	< .0001
Language use	5.48	11.02	0.497	= .627

Comparison of VOT Values for E-A and A-E Bilingual Speakers and A Monolingual Speakers Producing /k/ in English

Since the production of the voiceless plosive /k/ differs between the two groups of monolingual speakers, I compared the A-E and E-A bilingual speakers' English productions with those of E mono speakers. Figure 4.2 shows the numerical differences between E mono, A-E bilingual and E-A bilingual speakers when producing the English /k/. Again, I used mixed effects model to determine whether the monolingual English, E-A bilingual and A-E bilingual groups shown in Figure 4.2 differ statistically significantly in terms of VOT values for /k/. The statistical results shown in Table 4.12 demonstrate that E-A and A-E bilingual speakers differ from E mono speakers in terms of the VOT values of the voiceless stop /k/. Neither bilingual group showed target-like production in English, which means that the A-E bilinguals do not hit the L2 target VOT for the English /k/ and that the E-A bilinguals may show L1 attrition in their English /k/. Specifically, while the E-A bilinguals have significantly shorter VOTs for the English /k/ compared to the English monolinguals, both groups' VOT values are clearly within the long-lag range, so that it is not entirely clear if this is a case of attrition.

Table 4.11. VOT Value Comparison of E Mono and E-A and A-E Bilinguals for /k/ in English

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	111.16	5.58	41.29	19.921	< .0001
E mono vs A-E bilingual	30.41	7.52	28	-4.05	< .0001
E mono vs E-A bilingual	-19.451	7.52	28	-2.59	= .015

To explain the significant differences between the bilingual groups and the monolingual English speakers, I employed mixed effects model to determine whether aptitude and language use influence the VOT values of the English /k/ produced by A-E and

E-A bilingual speakers. Tables 4.13 and 4.14 summarise these statistical results, showing that bilingual speakers did not significantly differ in terms of the VOT values of the voiceless /k/ in English according to their aptitude or language use patterns.

Table 4.12. VOT Value of /k/ in English According to Aptitude Group

Fixed effects	**Estimate**	**Std. error**	***t*-value**	***p*-value**
Intercept	95.71	5.35	17.901	< .0001
Aptitude	-8.59	7.827	-1.097	= .293

Table 4.13. VOT Value of /k/ in English According to Language Use Group

Fixed effects	**Estimate**	**Std. error**	***t*-value**	***p*-value**
Intercept	91.9524	5.9737	15.393	< .0001
Language use	-0.4671	8.1798	-0.06	= .96

VOT Values for the Voiceless Plosive /t/ in Arabic and English
Comparison of Patterns for Monolingual Speakers

To determine whether E mono and A mono speakers differ when producing the voiceless /t/ in Arabic and English, I compared the VOT values for the two groups. Figure 4.3 shows the mean VOT values for the Arabic and English /t/ produced by bilinguals and A mono and E mono speakers. The VOT values for /t/ were higher for E mono than for A mono speakers, suggesting that the E mono group is likely to produce the /t/ vowel with stronger aspiration (longer lag) than is the A mono group.

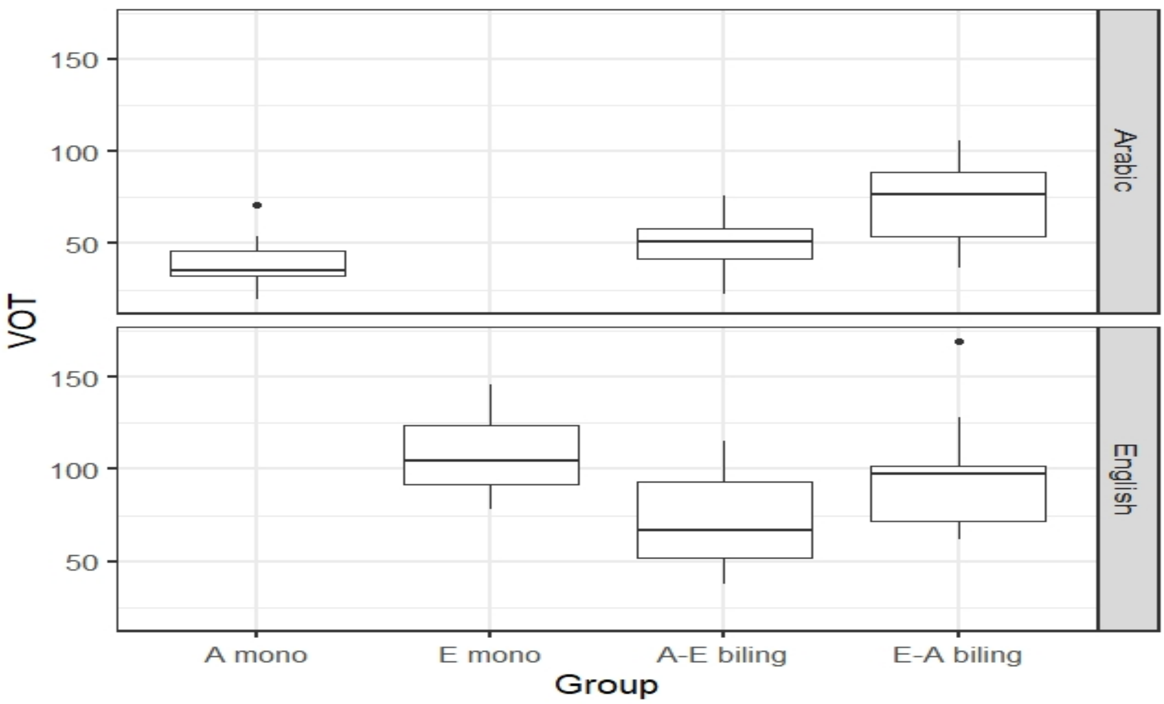


Figure 4.4. boxplots for the VOT value (in msec) of the Arabic and English /t/ produced by A-E and E-A bilinguals and the A mono and E mono groups

I used mixed effects model to determine whether the monolingual Arabic and English speakers shown in Figure 4.3 differ statistically significantly in terms of the VOT values for the voiceless stop /t/. The main effect for the speaker group variable in Table 4.15 represents a statistically significant difference in the VOT values between A mono and E mono speakers when producing /t/, with significantly stronger aspiration for E mono than for A mono speakers.

Table 4.14. VOT Value Comparison of A Mono and E Mono Groups for the Voiceless /t/

Fixed effects	**Estimate**	**Std. error**	**Df**	***t*-value**	***p*-value**
Intercept	38.43	4.73	28	8.131	< .0001
Speaker group	70.17	6.684	28	10.498	< .0001

4.7.4 Comparison of Patterns for Bilingual Speakers

Comparison between the VOT Values for E-A and A-E Bilingual Speakers and A Monolingual Speakers for /t/ in Arabic

Since the monolingual participants differ in their production of the /t/ voiceless plosive, it is also relevant to compare productions by A-E and E-A bilingual groups to those of monolingual Arabic speakers. Figure 4.3 shows the differences in the VOT values for /t/ produced by A mono, E-A bilingual and A-E bilingual speakers. I used mixed effects model to determine whether the E-A and A-E bilingual groups shown in Figure 4.3 differ in a statistically significant way. The results presented in Table 4.16 reveal that E-A bilingual speakers differ from A mono speakers in terms of the VOT value for the voiceless /k/, but no differences are observed between A-E bilingual and A mono speakers. These findings indicate that the E-A bilingual group does not achieve target-like production in terms of the VOT values for the voiceless stop /k/ in their L2 Arabic. In addition, the A-E bilinguals show no evidence of L1 attrition for the Arabic /k/ (even though the p-value obtained was close to reaching significance).

Table 4.15. VOT Value Comparison of A Mono and E-A and A-E Bilingual Groups for /t/ in Arabic

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	38.43	4.39	42	8.76	< .0001
A mono vs A-E bilingual	12.242	6.21	42	1.972	= .055
A mono vs E-A bilingual	32.911	6.21	42	5.302	< .0001

Because the Arabic monolingual speakers and the E-A bilingual speakers differ significantly in terms of the VOT values of /t/, I used mixed effects model to determine whether phonetic aptitude and language use influence these values within the E-A bilingual group. Tables 4.17 and 4.18 show the statistical results, revealing no significant difference between the aptitude groups or language use groups in terms of the VOT values of the voiceless Arabic /t/.

Table 4.16. VOT Value of /t/ in Arabic According to Aptitude Group

Fixed effects	Estimate	Std. error	*t*-value	*p*-value
Intercept	72.98	8.22	8.88	< .0001
Aptitude	-3.52	12.033	-0.292	= .775

Table 4.17. VOT Value of /t/ in Arabic According to Language Use Group

Fixed effects	Estimate	Std. error	*t*-value	*p*-value
Intercept	64.38	8.411	7.654	< .0001
Language use	13.06	11.52	1.134	= .277

Comparison between the VOT Values for E-A and A-E Bilingual Speakers and A Monolingual Speakers for /t/ in English

Next, because the productions of the voiceless plosive /t/ differ between the two monolingual groups, I compared the A-E and E-A bilingual groups' English productions with those of E mono speakers. Figure 4.3 illustrates the differences between the E mono, A-E bilingual and E-A bilingual speakers when producing the English /t/. Again, I used mixed effects model to determine whether the VOT values of /t/ in English differ significantly across the monolingual English and bilingual speakers. Table 4.16 presents the results. While the E-A bilingual group does not differ from the E mono group in a significant way, the A-E bilingual speakers differ from the E mono group in terms of the VOT values for the voiceless stop /t/. Once again, A-E bilingual speakers do not demonstrate target-like productions in English and there is no evidence for L1 attrition in the E-A bilingual group.

Table 4.18. VOT Value Comparison between E Mono and E-A and A-E Bilingual Groups for /t/ in English

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	108.592	6.42	41.141	16.922	< .0001
E mono vs A-E bilingual	-37.471	8.599	28	-4.358	< .0001
E mono vs E-A bilingual	-12.38	8.599	28	-1.439	= .161

Since significant differences exist between the A-E bilingual speakers and the monolingual English speakers, I employed mixed effects model to determine whether aptitude and language use influence the VOT values of /t/ in English produced by A-E bilingual speakers. The statistical results summarised in Tables 4.20 and 4.21 show that the bilingual speakers' aptitude and language use do not significantly influence the VOT values of the voiceless /t/ in English.

Table 4.19. VOT Value for /t/ in English According to Aptitude Group

Fixed effects	**Estimate**	**Std. error**	***t*-value**	***p*-value**
Intercept	98.14	10.443	9.397	< .0001
Aptitude	-4.112	15.29	-0.27	= .792

Table 4.20. VOT Value for /t/ in English According to Language Use Group

Fixed effects	**Estimate**	**Std. error**	***t*-value**	***p*-value**
Intercept	84.99	10.4	8.21	< .0001
Language use	21.05	14.18	1.49	= .161

4.7.5 VOT Values for the Voiceless Plosive /p/ in English

Comparison of VOT Values for E-A and A-E Bilingual Speakers with E Monolingual Speakers Producing /p/ in English

Since MSA does not have a /p/ sound, the E mono speakers could not be compared with the A mono group; I thus compared the E mono speakers directly against the A-E and E-A bilingual speakers. Figure 4.4 shows the mean VOT values for the English /p/ sound produced by E mono, A-E bilingual and E-A bilingual speakers. I used mixed effects model to identify any statistically significant differences between the bilingual groups and the monolingual English speakers in terms of the VOT values of the voiceless stop /p/, as shown in Figure 4.4. Table 4.22 presents the statistical results and shows that E-A and A-E bilingual speakers do not differ from E mono speakers in terms of the VOT values for the voiceless stop /p/. Both bilingual groups have achieved target-like English productions.

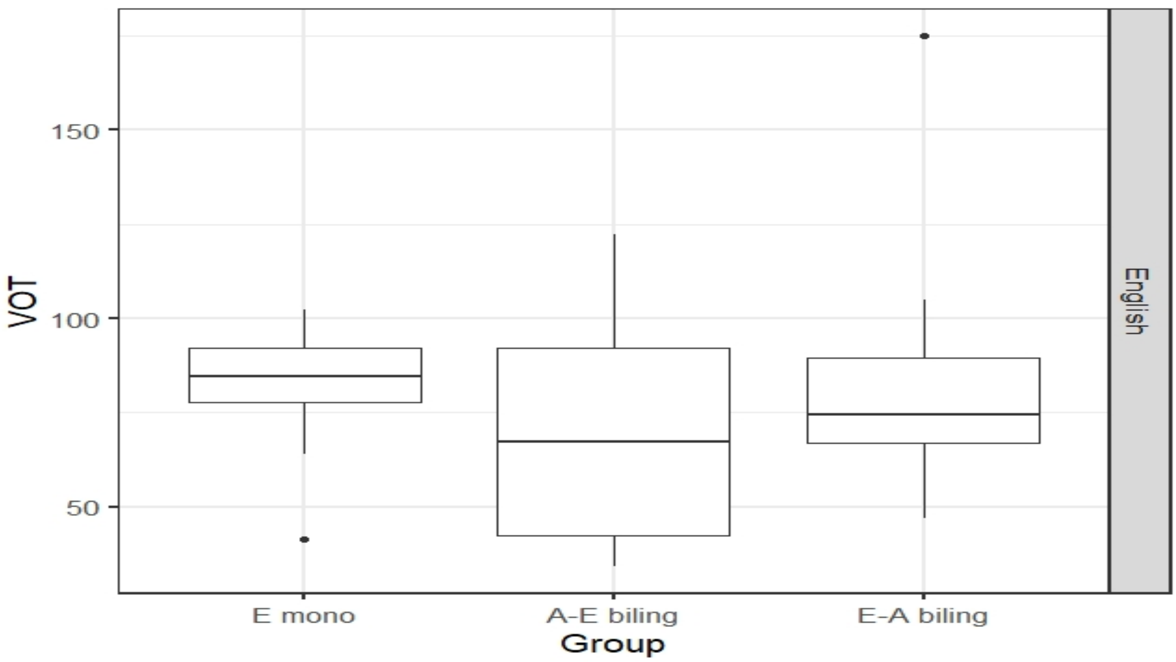


Figure 4.5. Boxplots for the VOT value (in msec) of the Arabic and English /p/ for A-E and E-A bilinguals and E mono groups.

Table 4.21. VOT Value Comparison of E mono and E-A and A-E Bilingual Groups for /p/ in English

Fixed effects	Estimate	Std. error	Df	*t*-value	*p*-value
Intercept	82.696	6.682	42	12.38	< .0001
A mono vs A-E bilingual	-14.12	9.45	42	-1.494	= .143
A mono vs E-A bilingual	-1.654	9.45	42	-0.18	= .862

Since the A-E and E-A bilingual groups do not differ in the VOT values of the /p/ production when compared to the E mono group, there is no need to consider whether language use or phonetic aptitude influences phoneme production by bilingual speakers.

4.8 Discussion

The current chapter examined the VOT values of voiceless plosives in L2 acquisition and L1 attrition for speakers of two typologically different languages, especially late E-A and A-E bilinguals. The chapter also examined whether bilinguals' phonetic aptitude and/or language use influenced how target-like were their productions of voiceless plosives in their L1 and L2.

As in Chapters 2 and 3, this chapter tested five hypotheses. In hypothesis one, I expected that there would be a difference in the pronunciation of voiceless plosives in Arabic and English. In particular, Hypothesis 1a suggested that the VOT of the voiceless stop of the Arabic monolingual speakers would be in the range of the English voiced (short-lag) stop. Hypothesis 1b proposed that the VOT of the voiceless stop of the English monolingual speakers would be strongly aspirated (long lag). A direct comparison of the VOT values of the voiceless stops /k/ and /t/ produced by the monolinguals confirmed Hypothesis 1: both /k/ and /t/ differed significantly across the two monolingual groups. In particular, the VOT values for Arabic /k/ and /t/ are in the short-lag range (similar to English voiced stops), whereas the VOT values for English /k/ and /t/ are strongly aspirated, i.e. in the long-lag range.

In Hypothesis 2, I predicted that A-E bilinguals would form new categories for /t/ and /k/ in English and Arabic and they would produce them with VOT values much the same as do native speakers of Arabic or English (Hypothesis 2a). I also predicted that E-A bilinguals would produce the VOTs for the voiceless Arabic plosives /t/ and /k/ in a manner similar to the VOTs of Arabic, and A-E bilinguals would produce the English plosives /t/ and /k/ in a manner similar to the VOTs of English. Finally, I expected that A-E bilinguals would acquire the /p/, such that their productions would be similar to those of native English speakers (Hypothesis 2b). Comparing the VOT values of bilinguals' L2 with that of monolinguals, I found that L2 learners of both Arabic and English failed to achieve the native patterns when pronouncing the voiceless plosives /k/ and /t/; thus, their results did not support Hypothesis 2a. By contrast, A-E bilinguals produced /p/ similarly to native English speakers, thus supporting Hypothesis 2b.

Hypothesis 3a predicted that A-E bilinguals would show L1 attrition and produce the VOTs for the voiceless Arabic plosives /t/ and /k/ in a manner similar to English, and these would therefore be longer than the VOTs of Arabic monolinguals. I also predicted that E-A bilinguals would produce the VOTs for the voiceless English plosives /t/ and /k/ in a manner similar to Arabic and that these would therefore be shorter than the VOTs of English monolinguals (Hypothesis 3b). Notably, I found attrition in the E-A bilinguals' production of the English voiceless stop /k/, but not in the /t/, and there was no such evidence for the A-E bilinguals. Thus, Hypothesis 3 was only partially confirmed.

I tentatively proposed that language use would affect L2 acquisition and L1 attrition (Hypothesis 4). Comparing language use in the bilinguals' productions suggested that there was no evidence for a link between language use and the bilinguals' productions, and therefore the current results did not confirm Hypothesis 4. Finally, based on Mennen's (2004) study, in which participants achieved native-like production either in both languages or in neither, I tentatively hypothesised that participants with high phonetic aptitude would have native-like VOT values in both their L1 and L2 (Hypothesis 5). However, Hypothesis 5 was not confirmed: similar to language use, phonetic aptitude did not affect L2 acquisition and L1 attrition. These results are discussed in detail in the following section.

4.8.1 Monolinguals

The current data from the monolingual English and Arabic groups confirmed Hypothesis 1 in that the VOT values of Arabic voiceless stops differed reliably from the VOT values of English voiceless stops. As expected, Arabic voiceless stops have a short-lag VOT pattern, and English have a long-lag VOT pattern. The /k/ and /t/ produced by the English monolingual group showed strong aspiration (long lag), which was not found for Arabic. The current findings for Arabic monolinguals confirm suggestions by the earlier literature, including Yeni-Komshian et al. (1977), Jesry (1996) and Docherty (1992). In particular, Yeni-Komshian et al. (1977) and Jesry (1996) proposed that the VOT value is between 20 to 30 msec for the Arabic /t/ and between 25 to 32 msec for the Arabic /k/. The current findings for English are in line with the VOT ranges measured in Docherty (1992), with a VOT range of 30-150 msec for /k/, 30-110 msec for /t/ and 10-80 msec for /p/. The findings from the current study of a short-lag pattern for Arabic and a long-lag aspirated pattern for English are compatible with this description. However, the findings from the Arabic monolinguals in this study are different from some previous studies. For example, Al Dahri's (2010) and Alghamdi's (2006) research suggested that the Arabic voiceless plosives are long-lag aspirated and do not fall into the voiced range of English stops. Thus, while Deuchar and Clark (1995), Yeni-Komshian et al. (1977) and Jesry (1996) proposed that Arabic voiceless stops /k/ and /t/ almost fall into the range of English phonemically voiced stops (short lag), Al Dahri's (2010) and Alghamdi's (2006) studies on the colloquial Saudi dialects Najdi and Gamdi differ from this pattern because the VOT in these dialects is about 30 msec longer than the VOT in other Arabic dialects.

In the following sections, I explain the bilinguals' results in light of the SLM hypothesis (Flege, 1995) and I discuss the language use and phonetic aptitude factors for both bilinguals L2 acquisition and L1 attrition.

4.8.2 Bilinguals: L2 Acquisition

Attempts have been made to describe L2 phonetic acquisition in relation to the L1 phonetic system (SLM, Flege, 1995). In the current chapter, three different acquisition patterns have been observed: complete L1-to-L2 transfer, equivalence classification and native-like acquisition. I found that both groups of highly advanced late L2 bilinguals (A-E and E-A bilinguals) showed L1-to-L2 transfer. The complete L1-to-L2 transfer was observed in the A-E bilinguals for the English voiceless stops /k/ and /t/ and in the E-A bilinguals for the Arabic /t/ and /k/, thus failing to support Hypothesis 2a. In particular, I found statistically significant differences between the monolinguals' and the bilinguals' L2 speech, suggesting that the bilinguals in this chapter did not manage to achieve native like acquisition of the L2. This means their results did not support the SLM, which predicts that bilinguals establish a new category for the L2 sounds and produce the L2 sounds differently from the L1 sounds and are potentially native-like in their L2 productions. More precisely, I found evidence that the bilingual productions of the /t/ and /k/ fell in between the two languages' monolingual norms, and were thus more similar to bilinguals' L1 productions than native-like L2 acquisition would be. These results are in line with previous studies showing transfer from L1 to L2 (Flege, 1987; Flege & Port, 1981). However, unlike the participants in the current study, the participants in Flege (1987) and Flege and Port (1981) were learners of English and not highly proficient L2 learners. Therefore, as the current results found L1-to-L2 transfer even for highly proficient speakers who had lived in the host country for a long time, suggesting that native-like VOT values may be difficult to acquire for late learners, even highly proficient ones.

The results of this chapter showed that A-E bilinguals acquired the /p/ like native English speakers, which suggests that the A-E bilinguals in this study managed to approximate the VOT value of the voiceless /p/ in their L2, thus supporting Hypothesis 2b. It is perhaps unsurprising that highly proficient bilinguals, especially those who have been in the host country (where the L2 is spoken) for a long time, approximate native norms for sounds that have no close equivalents in their native language. My results are in line with previous studies (Simon, 2009; Simon & Leuschner, 2010; Schmid et al., 2014), especially those of Schmid et al. (2014), even though their study of advanced Dutch-English bilingual participants was conducted on voiced and voiceless plosives and they observed native-like acquisition for L2 short-lag voiced stops. The A-E bilinguals' results for /p/ support the SLM theory, which predicts that learners establish new category for phones that are new, and no new category for phones that are similar. So the SLM predicts native-like productions for /p/, but not (or only later) for /t/ and /k/.

The data in this chapter relating to L2 acquisition of voiceless stops did not support Hypothesis 2a, which suggested that both groups of bilinguals would form new L2 categories for /t/ and /k/. However, the data did support Hypothesis 2b, which suggested that the A-E bilinguals would create a new category for the English /p/. While neither group of bilinguals achieved proper L2 pronunciation, the A-E bilinguals achieved native-like pronunciation in the production of the English voiceless /p/.

The findings of the present chapter also failed to confirm Hypotheses 4 and 5 because neither language use nor phonetic aptitude showed any significant influence on participants' voiceless plosive production; hence, there was no evidence that differences in production between participants with high and low L2 use or high and low aptitude related to the nativeness of their production. This finding contradicts previous research (e.g. Jilka et al., 2008; Hopp & Schmid, 2010). Jilka et al.'s (2008) work on sounds revealed that pronunciation of German-English bilinguals was influenced by the participants' phonetic aptitude. This difference in results may be because Jilka et al. (2008) had a very large sample size, whereas this study was carried out on a significantly smaller scale. Moreover, the tests used in this study to measure the phonetic aptitude are not exactly the same as those used by Jilka et al. (2008). However, these results are in line with the results in Chapter Three, suggesting that language use and phonetic aptitude – at least in the current study – do not affect VOT values for voiceless plosives (this chapter) or vowel formants for shared vowels (Chapter Three).

4.8.3 Bilinguals: L1 Attrition

The data in this study did not or only partially confirmed Hypothesis 3 in that I possibly found evidence for L1 attrition, but only among the L1 English bilinguals' productions of the voiceless stop /k/, and not among the L1 Arabic bilinguals. Specifically, the E-A bilinguals produced the English voiceless stop /k/ with significantly shorter VOT values than did the English monolinguals. Therefore, E-A bilinguals produced /k/ in both languages, Arabic and English, with longer VOT values than the Arabic norm, but not as long as the aspirated stops of English. Assimilatory patterns with compromise VOT values in this way have been confirmed in a number of previous studies on L1 attrition (Flege, 1987; Major, 1992; Mayr et al., 2012). The L1 VOT drift in the English voiceless plosive /k/ is also in line with the SLM because it assumes flexibility in the phonetic categories of the L1 (Flege, 1995). However, the English /k/ values for E-A bilinguals, while statistically significantly shorter than for E monolinguals, are still clearly in the long-lag range. As mentioned above, it is therefore not entirely clear if this should be considered to be a case of L1 attrition. Perception studies would need to determine if this is indeed a case of attrition that monolingual native speakers can perceive or if it is merely variation within the native range.

Overall, the bilinguals' results showed no attrition in their L1 voiceless plosives, except for possibly the /k/ produced by the E-A bilinguals. It is not immediately clear why this is the only sound that may show attrition by the E-A bilinguals; there is no obvious reason or explanation for why /k/ should be affected but not /t/.

Similar to the findings regarding vowels, the voiceless stop attrition data did not support Hypotheses 4 or 5. Specifically, the A-E bilinguals showed no L1 attrition, and neither L2 language use nor phonetic aptitude had an impact on the A-E and E-A bilinguals' production in a way that one group (high or low aptitude and more or less L1/L2 use) could be considered to be more native-like. The results that language use in the current study had no effect on the bilinguals is quite surprising and different from the study of Sancier and Fowler (1997). The current study differs from Sancier and Fowler's study because Sancier and Fowler had only one participant, who was specifically selected because other speakers had observed L1 attrition in her speech. Thus, the present study and Sancier and Fowler were asking different research questions: I wanted to know if there would be L1 attrition in terms of VOT in two groups of bilingual speakers, while Sancier and Fowler had one bilingual who, according to other people's observations, was showing L1 attrition; and they wanted to know if (in a speaker with observed attrition) the language of the surroundings would module VOT values.

4.8.4 Conclusion

The current chapter confirmed that the VOT values of voiceless stops in Arabic and English differ reliably. The results that the E-A bilinguals' English /k/ differs reliably from that of monolingual English speakers is broadly in line with the SLM's assumed flexibility of phonetic categories in the L1. In addition, the bilinguals' results of the voiceless plosives /t/ and /k/ showed L1-to-L2 transfer. Moreover, the results for A-E bilinguals' acquisition of /p/, but not /t/ or /k/, in the present chapter support the SLM theory, which predicts that speakers establish new categories in the L2 for L2 sounds that differ sufficiently from L1 sounds. Moreover, language use and phonetic aptitude did not contribute to the acquisition of voiceless plosives in either Arabic or English.

Chapter Five: General Discussion and Conclusions

5.1 Introduction

This project examined three different aspects of the productions of late E-A and A-E bilinguals. Specifically, this included the intonation patterns of *wh*-words in Chapter Two, the shared vowels' formants in Chapter Three and the VOT of voiceless plosives in Chapter Four. In addition, the current project examined whether language use and phonetic talent may affect L1 attrition and L2 acquisition. The literature reviews in Chapters Two, Three and Four show that a number of studies suggest a bi-directional influence of L1 and L2 on the production of segmentals (Flege & Hillenbrand, 1984; Flege, 1987; Major, 1992; Flege, 1994, 1995; Chang, 2012; Mayr et al., 2012; de Leeuw et al., 2013) and suprasegmentals (Willems, 1982; McGory, 1997; Ueyama, 1997; Mennen, 1999; Guion et al., 2000; Jilka, 2000; Mennen, 2004; de Leeuw, 2009). In addition, it shows that there are some factors that seem likely to have affected L1 attrition and L2 acquisition (Leather, 1987; Flege & Fletcher, 1992; Guion et al., 2000; de Leeuw, 2009; Mayr et al., 2012; Stoehr et al., 2017). Overall, it appears that certain kinds of language use have an influence on L1 attrition and L2 acquisition, but it is not clear which aspects of pronunciation may be affected by language use. In addition, Mennen's (2004) study finds variability across individuals and this could be due to phonetic talent or due to a variety of other factors that may be related to native-likeness in both the L1 and the L2. The empirical evidence from the current study revealed that there is a bi-directional influence between the L1 and L2 at both the segmental and suprasegmental levels, and it also suggested that attrition may not affect all areas of pronunciation equally.

To achieve the main aims of the project, late highly fluent A-E and E-A bilinguals and Arabic and English monolinguals completed three different production tasks: reading a list of 12 *wh*-questions and eight yes/no questions (for the prosody analysis), reading a list of words and sentences that contained different vowels (for the formant analysis) and narrating short cartoons (for the VOT analysis). The bilinguals also provided proficiency tests for Arabic and English and a phonetic aptitude test. Given the research aims, the hypotheses were based on the SLM and previous research on L1 attrition and L2 acquisition regarding the production of segmentals and suprasegmentals in English and other languages.

This chapter will start by presenting an overview of the key findings in section 5.2 and a general discussion of the results in section 5.3. Section 5.4 considers the project's implications and Section 5.5 presents its limitations and recommendations for future research. The conclusions are presented in section 5.6.

5.2 Overview of the Key Findings

The present project is the first to examine L1 attrition and L2 acquisition in the speech of late, highly proficient L2 bilinguals of Arabic and English, which are two typologically different languages. The main purpose of the project was to explore L2 acquisition and L1 attrition in the production of *wh*-words, of formants of shared vowels and of voiceless stops in late bilinguals. In addition, the aim was to examine whether late L2 learners' language use and phonetic aptitude affect how target-like learners' productions of segmentals (vowel formants and the VOT of voiceless plosives) and suprasegmentals (prosody of *wh*-words) are in both Arabic and English. This was measured by comparing the speech of the bilinguals who had lived in an L2 environment (for about 20 years) with that of Arabic and English monolinguals. Three different areas of pronunciation were examined: Study I focused on *wh*-word production, Study II examined formants of shared vowels and Study III examined the VOT in plosive productions. In this section, I will discuss the findings from the project as a whole. First, I will discuss the monolinguals of both languages, the bilinguals regarding L2 acquisition and the bilinguals regarding L1 attrition. Then, I will discuss the role of language use and phonetic aptitude in L2 acquisition and L1 attrition.

5.2.1 Monolinguals

The hypotheses regarding monolinguals were formulated on the basis of previous studies which found that Arabic and English differ in the production of *wh*-words (Defense Language Institute, 1974; El Zarka, 1997; El Zarka, 2017; Bartels, 2014), *shared* vowels (Deterding, 2006; Alotaibi & Hussain, 2009) and the *VOT* of voiceless stops (Yeni-Komshian et al., 1977; Docherty, 1992; Jesry, 1996). The current study supports these previous findings. The *wh*-words produced by the group of Arabic monolinguals showed a steep rise-fall pattern, which was not found for English, supporting previous work showing that the typical pitch pattern for Arabic *wh*-words is different from the typical pitch pattern for English *wh*-words (Defense Language Institute, 1974; El Zarka, 1997; Bartels, 2014).

Furthermore, monolingual vowel productions suggested differences in the productions of the vowels shared between Arabic and English. The Arabic monolinguals produced /iː/, /ɪ/, /uː/ and /ʊ/ with a lower tongue position than the English monolinguals. English monolinguals produced the long /ɑ:/ and /a/ with similar tongue position. The Arabic monolinguals produced the /a/, /ɪ/, /i:/ /u:/ vowels with less fronted compared to the English monolinguals. Moreover, both monolingual groups produced /ɑ:/ similarly in terms of tongue frontedness.

The last key finding is that the monolinguals in this study suggest that the VOT values of Arabic voiceless stops differed reliably from the VOT values of English voiceless stops. Previous studies have suggested that voiceless stops in English have a long lag VOT pattern and voiceless stops in Arabic have a short-lag VOT pattern. The /k/ and /t/ produced by the English monolingual group was strongly aspirated (long lag), which was not found for Arabic monolinguals. The current study's findings of VOT values for Arabic and English voiceless stops support previous findings (Yeni-Komshian et al., 1977; Docherty, 1992; Jesry, 1996).

5.2.2 L2 Acquisition

The hypotheses regarding the bilinguals' L2 acquisition of *wh*-words were formulated based on previous research, which expected that the highly proficient E-A and A-E bilinguals would approximate the L1 pattern in their L2. The major finding was that the L2 acquisition of the *wh*-words indicated that highly advanced, late L2 bilinguals approximated native productions when producing *wh*-words. More specifically, despite the fact that the bilinguals' L2 *wh*-word patterns differed from those of monolinguals, there was no evidence to suggest that the differences found were influenced by the bilinguals' L1. It is not surprising that highly proficient bilinguals approximate native norms, particularly those who have been living in the L2 environment for a long time. However, the bilinguals' results were somewhat surprising, as they were found to be contrary to cross-linguistic interference predictions. The *wh*-word results of the bilinguals in their L2 confirmed LILt, which holds that, as learners increase their experience of the L2, their intonation parameters will more closely approximate the norms of the L2.

The hypotheses regarding the bilinguals' L2 acquisition of vowels were formulated based on previous research and the SLM and PAM-L2. The SLM and PAM-L2 predicted that the participants may not form a new vowel category for L2 vowels that are similar to L1 vowels. I therefore expected that the bilingual participants would pronounce the L2 vowels as they do in their L1. Thus, bilinguals' L2 vowels are expected to differ from those of the monolingual participants. The results for A-E and the E-A bilinguals showed that these two groups achieved native English-like and native Arabic-like production of some of the shared vowels. This result is not what I predicted based on the SLM and PAM-L2, but it is nevertheless compatible with these two theories as they do allow for native-like production of similar vowels and merely assume that this is less likely to occur than in the case of L2 vowels that are sufficiently distinct from L2 vowels so that learners readily create a new category. However, it is perhaps not so surprising that bilinguals who are highly proficient in their L2 approximate native norms, particularly those who have been in the L2 environment for a long time. However, the study found that E-A bilinguals did not achieve native-like pronunciation of Arabic vowels. Thus, the vowel findings for the E-A bilinguals, but not the A-E bilinguals, align with those of previous studies that support the SLM theory, such as Munro et al. (1996), Flege (1987) and Flege and Hillenbrand (1984). In addition, the findings from the study displayed a significant effect of speaking condition for some vowels and showed that the vowel productions were significantly less peripheral in a carrier phrase than for words in isolation, particularly in the production of the Arabic vowel /iː/. This finding is in line with Deterding's (1997) study.

The hypotheses regarding the bilinguals' L2 acquisition of the VOT of plosives were formulated based on previous research and the SLM. In terms of the SLM, bilinguals would establish a new category for the L2 sounds and produce the L2 sounds differently from the L1 and potentially be native-like in their L2 productions (Flege, 1995). In the analyses of the VOT, only the A-E bilinguals established a new category and achieved native-like productions in the L2. The English-Arabic bilinguals' results show an effect of the first language (English) on the L2 (Arabic), with their Arabic voiceless stop /t/ realised with strong aspiration. Meanwhile, the E-A bilinguals did not establish a new category for the Arabic plosives. Overall, support for the SLM from the current study is mixed with the findings of the A-E bilinguals' vowel data and the E-A bilinguals' VOT data supporting the SLM.

5.2.3 L1 Attrition

English-Arabic bilinguals showed attrition in their English *wh*-word production, suggesting an influence of the steep rise-fall pattern of Arabic (their L2). L1 attrition also appeared in Study III (Chapter Four) in E-A bilinguals, displayed in the production of the English voiceless plosive /k/, suggesting an influence of their L2 Arabic's short-lag categories. In Study II (Chapter Three), both E-A bilinguals and A-E bilinguals also showed attrition in their L1, as they produced some of their Arabic and English vowels differently from Arabic and English monolinguals. These findings suggest that attrition may not affect all areas of pronunciation equally, and are consistent with those of Mayr et al. (2012). In addition, the findings are consistent with Hopp and Schmid's (2013) study, which concluded that "acquiring a language from birth is not sufficient for ensuring nativelikeness in bilingual speech production" (p.388).

Generally, the present project suggests that living in the L2 environment for a long time affects the production of the L1 of late A-E and E-A bilinguals. In this regard, even though this project looked at the data as a group and not individually, the late bilinguals' findings as a group are different from results found for some individuals in previous research, who achieved native-like production of their L2 and L1 for both languages (Major, 1992; Mennen, 2004). The particular factors that caused the L1 attrition in the bilinguals' production are not known. However, language use, but not phonetic aptitude (or at least to a much lesser extent), modulates how closely bilingual participants approximate native patterns in the production of *wh*-words.

However, not all the examined production areas of the late A-E and E-A bilinguals were liable to L1 attrition. For example, the English /t/ was produced by E-A bilinguals in a native manner, the Arabic /ɑː/ and /a/ were nativelike in A-E bilinguals productions, and the English ɑː/, /a/ and /ʊ/ were native-like in the productions of E-A bilinguals. In line with recent proposals, the current study found bidirectional influences such that L1 and L2 interact with one another. That is, not only does the L1 affect the L2, the L2 also has an influence on the L1. In addition, the results of the bilinguals in the present project suggest that some areas of pronunciation may be more vulnerable to attrition than others and that some areas of pronunciation may also be more amenable to achieving native-likeness than others.

5.3 General Discussion

In the previous section, the results of Studies I, II and III were summarised with respect to the monolinguals' results and the bilinguals' results in L2 acquisition and L1 attrition. This section provides a holistic judgement of the findings of the three studies in the areas of pronunciation, the relationship between L1 attrition and L2 acquisition, the L2 speech learning models, and the role of phonetic aptitude and language use.

First, the results of the bilinguals suggested that some areas of pronunciation may be more vulnerable to attrition than others and also more amenable for achieving native-likeness than others. In all three studies, L1 attrition was observed. In particular, L1 attrition appeared in the production of *wh*-words for the L1 English bilinguals but not in the L1 Arabic bilinguals. In addition, the L1 English bilinguals may have shown attrition in the production of the voiceless plosive /k/, but this was not the case for the L1 Arabic bilinguals in the L1 Arabic. The L1 Arabic bilinguals showed attrition in their L1 vowels /ɪ/, /iː/, /ʊ/ and /uː/, whereas the L1 English bilinguals they showed attrition in their L1 vowels /ɪ/, /iː/, and /uː/. There are several possible reasons for these discordant results of the E-A and A-E bilinguals in L1 attrition.

In particular, the results for the prosody of *wh*-words are compatible with Eckman's (1987, p.1) MDH, which states that "when two languages differ, marked structures are more difficult to acquire than unmarked structures". English sentence prosody is considered to be more marked than Arabic sentence prosody (Zerbian, 2015) and is therefore predicted to be more difficult to learn than Arabic sentence prosody. Therefore, it is possible that the English prosody of *wh*-words is harder to acquire than the Arabic prosody of *wh*-words. If a particular characteristic is harder to learn in the L2, then it typically takes longer to acquire and may never be fully mastered. One possibility is that aspects that are harder to acquire are also harder to lose, as increased effort has initially been put into acquiring these aspects. This might possibly explain why L1 attrition was found in the prosody of the E-A bilinguals, but not the A-E bilinguals. However, other studies have shown that marked patterns can be subject to L1 attrition, for example, in Mayr et al. (2012), MZ showed attrition in the marked Dutch front rounded vowels. Furthermore, as far as I am aware, there are no previous claims or demonstrations that markedness would affect L1 attrition. Consequently, the current study cannot draw any conclusions regarding a possible relationship between markedness and L1 attrition.

Successful L2 acquisition was observed when both groups of bilinguals approximated native productions of *wh*-words in their L2. However, their L2 productions of the segmental voiceless stops /k/ and /t/ differed from L1 productions. Additionally, the A-E bilinguals succeeded in correctly producing the English vowels and the English voiceless stop /p/, whereas the E-A bilingual did not correctly produce the Arabic vowels and Arabic voiceless stops. Based on these findings, A-E and E-A bilinguals can maintain native-like abilities in their L1 and L2 when producing *wh*-words (suprasegmental).

The results confirmed the idea that some areas of pronunciation may be more vulnerable to attrition than others and that some areas of pronunciation may also be more likely to achieve native-likeness than others.

The current project partially supported the predictions of the SLM theory and observed three different acquisition patterns: complete L1-to-L2 transfer, equivalence classification and native-like acquisition. Generally, the present work found some evidence for the SLM, particularly in the results of the vowels for E-A bilinguals in Chapter Three (Study II). Specifically, the results suggest that similar (but not identical) sounds in learners' L1 and L2 are difficult to learn, and that L2 speakers do not form a native-like L2 category for similar L2 sounds. The SLM also proposes that vowel categories for similar vowels in L1 and L2 are more likely to merge, while sounds that are not similar (different) are more likely to retain their separate categories: "By hypothesis, category formation will be blocked if instances of an L2 speech category continue to be identified as instances of an L1 category. The SLM predicts that in such cases, a 'merged' category will develop over time that subsumes the phonetic properties of the perceptually linked L1 and L2 speech sounds" (Flege et al., 2003: p.469).

In addition, study I on the prosody of *wh*-words is in line with LILt, as I have found that the E-A and A-E bilinguals approximate the norms of the L2. With regards to L1 attrition, the results from the E-A bilinguals, but not the A-E bilinguals, are in line with LILt, which allows for the L2 to influence the L2. Specifically, the E-A bilinguals showed attrition in their L1 *wh*-words, thus revealing plasticity in the L1.

In terms of acquisition, in the analysis of vowels, it appears that the results for the E-A bilinguals and the A-E bilinguals align with those of previous studies that support the SLM theory (e.g. Flege & Hillenbrand, 1984; Flege, 1987; Munro et al., 1996). In addition, the results of A-E bilinguals in the production of the English /p/ support the SLM. Therefore, the data support the SLM in some areas at the segmental level.

The previous studies discussed in the current project, which either directly or indirectly examined second-language acquisition or first-language attrition in the phonetic domain at the suprasegmental or segmental level, focused on one level (Weinreich 1954; Flege, 1994, 1995; McGory, 1997; Ueyama, 1997; Mennen, 1999; Jilka, 2000; Chang, 2012; Mayr et al., 2012; de Leeuw et al., 2013). The present thesis carries new information because it investigated the segmental level and the suprasegmental level for the same speakers in their L2 and L1. In addition, this study examined the relationship between L1 attrition and L2 acquisition within the same speakers. The results of the vowel data of A-E bilinguals and the prosody of *wh*-words of E-A bilinguals revealed that native-like L2 acquisition occurred along with L1 attrition, which means that sounds that have been acquired well in the L2 are susceptible to attrition in their L1. However, this is not the case for the present data of *VOT* values of voiceless plosives, as the E-A bilinguals did not achieve the Arabic voiceless stops but instead showed attrition in their L1 /k/. The place of articulation also seems likely to play a role in L1 attrition.

The variables that were considered to play a role in L1 attrition and L2 acquisition in the speech production of bilinguals were language use and phonetic aptitude. From the L1 attrition perspective, no evidence suggested that language use or phonetic aptitude influenced L1 attrition in the three studies; in other words, there was no effect of language use on the process of L1 attrition in the three studies. The main limitation was the method used to measure or quantify language use in the present project, and possibly any investigation into the linguistics of bilinguals. Specifically, there are many ways to collect information about language use and quantify this information. Thus, it is possible that language use has little effect on L1 attrition and L2 acquisition in the phonetic domain, or that the current project merely measured aspects of language use that do not affect attrition or acquisition and did not capture aspects of language use that may relate to attrition and acquisition. The present study revealed that frequent L1 use by various speakers with different phonetic aptitudes plays a small role in avoiding L1 attrition in the phonetic domain. Previous literature has discussed the difficulty of measuring aptitude; for example, de Leeuw (2008) decided not to investigate the influence of phonetic aptitude on the L1 attrition results because there was no established instrument available to measure this variable (Meara, 2006). Hence, the current project investigated the impact of phonetic aptitude on L1 attrition outcomes by creating an aptitude test using samples from the MLAT (Carroll & Span, 1959) and the PLAB (Pimsleur, 1968), which are available online. In conclusion, phonetic aptitude is not a clearly understood concept, and there is still no standardised tool available to measure it.

The present project, particularly Study I (Chapter Two), found that aptitude plays a small role in L2 acquisition, while some evidence was found of language use affecting pronunciation. More specifically, participants with more L2 language use showed more native-like production of their L2 than participants with less L2 language use. These results are in line with previous research by Purcell and Suter (1980), as well as with Hopp and Schmid's (2010) work. Despite the fact that phonetic aptitude did influence the L2 participants' productions, there was no evidence that the differences in production between participants with both high and low aptitudes were related to how native-like the productions were. In other words, the prosody study (Study I) revealed that neither group of bilinguals appeared more or less native-like in their L2.

The findings offer two broad conclusions: language use affected the prosody of L2 acquisition in the current study, but not any other aspects (segmentals, namely vowels and VOT of voiceless plosives). Second, L1 use and phonetic aptitude may not be essential for avoiding L1 attrition in the phonetic domain. I only found effects of talent when it comes to prosody, but not when it comes to segments. There are no clear reasons to suggest why these factors affect L2 prosody acquisition, but not the other phonetic levels explored in this thesis. One possible explanation is that second language learning typically does not involve instruction in prosody, but learners may systematically have learned individual segments during their L2 acquisition in a formal, i.e. school, context. It is thus possible that L2 language use has a larger influence on aspects of pronunciation that are not systematically learned in a more formal and systematic context. Phonetic aptitude and language use will be key themes for future research to identify conditions that lead to L2 success and L1 maintenance.

5.4 Implications of the Study

The present project has a number of implications for L1 attrition and L2 acquisition in the phonetic domain. It looked at segmental and suprasegmental aspects of the same late A-E and E-A bilinguals (in both their L1 and L2). Having the data from the same speakers in their second language offers the chance to test the data directly for any asymmetry in terms of deviance from the native norm in the L1 and the L2. In addition, Arabic and English are typologically different languages. To my knowledge, the present study is the first to directly compare the production of bilinguals in terms of L1 attrition and L2 acquisition in two typologically different languages and to examine whether language use and phonetic aptitude play a role in their L1 attrition and L2 acquisition. The vast majority of linguistic and phonetic research has investigated Indo-European languages, with a focus on Germanic languages. As Favier et al., (2019) have argued, linguistic research, especially research that is related to any aspect of cognition, would benefit from investigating typologically different languages. Restricting the range of languages may restrict the phenomena that research uncovers, which may limit the development of theories of first language attrition and second language acquisition or which may lead one to draw conclusions based on incomplete data.

The current study offers a better understanding of the nature of how sound systems interact in bilinguals who speak two typologically different languages. For example, the current work gives us a better understanding of the intonations of *wh*-words in Arabic and English, since very few studies on the intonation of *wh*-questions have focused on the *wh*-word; most studies have instead concentrated on utterance final intonation in *wh*-questions rather than on the actual *wh*-words themselves.

The current project examined whether aptitude and language use play a role in L2 acquisition and L1 attrition in the phonetic domain. This objective contributes to the growing number of studies exploring these factors in L2 language acquisition and L1 attrition and offers a better understanding of these factors. In addition, the present work contributes to the area of methodology. It is one of the few studies to use phonetic aptitude tests to measure aptitude directly to ascertain whether aptitude plays a role in L2 language acquisition and L1 attrition of segmental and suprasegmental productions of A-E and E A bilinguals. However, while using aptitude tests is a step in the right direction, it needs to be said that the concepts of aptitude and phonetic aptitude are still somewhat mysterious and a better definition of these concepts as well as research on how to best measure different aspects of aptitude are needed.

Maybe even more important than the theoretical contribution to the rather new area of the relationship between L1 attrition and L2 acquisition is the relevance that the present project offers to immigrants who have moved to a new country after the age of puberty or in adulthood. The United Nations estimates that 258 million people live outside of their country of birth (United Nations, 2017). Almost all of these people have acquired a new language as adults in their host country. Those people (immigrants) who wonder whether the pronunciation of their L1, as acquired from birth, has been influenced by the acquisition of their L2, may possibly find this research relevant.

In addition, the present project provides a better understanding of the factors involved in L1 attrition and L2 acquisition; it illuminates the processes of language loss and learning within the brain and could contribute to the development of pedagogies that can succeed in optimising L2 acquisition without concomitant L1 attrition. It may be possible, for example, to predict which phonological difficulties might arise in particular groups of learners and then to design appropriate instruction methods to remedy these difficulties in the classroom (Eckman, 1981).

5.5 Limitations of the Study and Recommendations for Future Research

The present project has succeeded in achieving its aim and has contributed to theoretical and methodological knowledge in the areas of L1 attrition and L2 acquisition in the phonetic domain for A-E and E-A late bilinguals in particular. Nevertheless, the study encountered a few possible limitations that will need to be addressed in the future by anyone who replicates the investigation or continues on from it. First, the genders were not equal or well-balanced across the four groups; there were more females than males in each group. Second, due to the sheer volume of data, a full-fledged manual or hand-corrected ToBI analysis of *wh*-words in Chapter Two was not carried out. A ToBI analysis may give a clearer picture of the pitch accent distribution on *wh*-words in English and Arabic; future studies will need to confirm the current findings regarding the prosody of *wh*-words in Arabic and English. Third, the production data from suprasegmental tasks in the current project were carefully designed by the researcher and were read many times by native and non-native speakers, but it might have been more natural to use data from authentic conversational materials. Data obtained using this type of source could provide a different picture of the participants' L1 and L2 abilities than what is shown here.

A further limitation is the method of sampling and matching of accents. In the current thesis, I had very narrowly-defined participant groups. For example, the bilingual participants had to be very proficient in both languages, and needed to have lived in the L2 environment for a long time. I used snowball method because of these very strict inclusion criteria. This method allowed me to more easily find people who fulfilled the inclusion criteria. However, the snowball method also influenced the kinds of participants who were selected. Specifically, as each participant suggested another participant, as a group, participants are not very socially diverse and likely to come from similar social circles, with some participants knowing each other. While this may not necessarily have influenced the results, it needs to be kept in mind when interpreting the results

There are different accents and dialects for both English and Arabic, which may differ substantially. Matching on accents is therefore challenging. My main concern was that the accents are matched well across participant groups, not within groups. In other words, while I allowed for participants from different regions within groups, I tried to ensure that regions were matched as well as possible across groups, keeping in mind the strict inclusion criteria. Thus, within groups participants were from different regions in Saudi Arabia (Makkah, Jeddah, Riyadh and Abha) and from different regions of the UK (Sheffield, Chester and London). But across groups, region was matched as well as was possible within the constraints of the inclusion criteria. For example, if the Arabic monolingual group included a participant from Makkah, I also tried to have a participant from Makkah or a region with a very similar dialect in the A-E bilingual group and a participant who lived in Makkah or a region with a very similar dialect in the E-A bilingual group. And I tried to avoid a situation where one group, for example, includes speakers from Makkah and another group speakers from Abha. That said, it does need to be acknowledged that speakers overall come from different dialect regions and that this may have introduced variability in the data that could have influenced the results. To sum up, the main concern was to match the participants across groups and matching within groups was secondary, and it was just not possible to find enough people to match both across groups and within groups very narrowly.

Also, this study did not fully determine how attrition and acquisition may relate to each other. The present research compared both bilingual groups' Arabic with Arabic monolingual speakers and both bilingual groups' English with English monolingual speakers. However, it did not compare the bilinguals directly in their L1 and L2 as this was beyond the scope of the current study. My results therefore speak to L1 attrition and L2 acquisition within the same individuals (in that I examined how they compared with monolingual native speakers of both of their languages). However, my results do not speak to whether there are any differences between L2 English vs. L1-attrited English or L2 Arabic vs. L1-attrited Arabic because I have not compared bilingual groups directly with each other. In addition, my results do not speak to whether bilinguals show distinct (or merged) patterns in their two languages because I have not compared the English and Arabic of the A-E bilinguals or the English and Arabic of the E-A bilinguals directly. This means that my results speak to the relation between L1 attrition and L2 acquistion within individuals, but not within languages. This will be reserved for future analyses. Therefore, future analyses that compare the bilinguals' patterns in the L1 and L2 would make a valuable contribution to understanding the relationship between L1 attrition and L2 acquisition. Additional analyses comparing the bilinguals' patterns across both their languages would clarify this relationship.

Furthermore, I have used a rather lax definition of phonetic aptitude when selecting the tests to assess phonetic aptitude. Specifically, I selected three tests to assess phonetic aptitude that involved knowledge and skills about segments or suprasegmentals, but that were not restricted to such knowledge and may have involved knowledge and skills unrelated to phonetic aspects of language. The test that I used that possibly most closely taps into what one may call phonetic aptitude is the 'sound discrimination test'. Thus, an alternative for measuring phonetic aptitude in the current study would have been to only use the 'sound discrimination test' as a measure of phonetic aptitude. The results from this test may have more narrowly and more accurately captured participants' phonetic aptitude.

Moreover, the current project mainly examined the production at the level of segmentals and suprasegmentals, and the results from the A-E and E-A bilinguals in this project suggested that there is a bidirectional influence of L1 and L2 sounds. Therefore, a future perception study on L1 attrition and L2 acquisition would be valuable to the field of phonetics and could improve the understanding of the nature of interactions in the sound systems of bilinguals. Such a study should examine whether the interplay of two typologically different language groups affects the perceived pronunciation accuracy in different ways. Finally, additional studies on the segmental and suprasegmental production of L1 attrition and L2 acquisition would improve the knowledge in this field and narrow the gap in the literature, completing the picture of intonation, formant and *VOT* value variation in English and Arabic.

5.6 Conclusions

This study carried out a direct comparison of the production of segmentals (formants of shared vowels and *VOT* values of voiceless plosives) and suprasegmentals (prosody of *wh*-words) in late L1 attrition and L2 acquisition of Arabic-English and English-Arabic bilinguals. A direct comparison between L1 attrition and L2 acquisition in the same speaker was presented for both their L1 and L2. The findings from this project reveal that the first language can be susceptible to attrition in the domain of phonetics. At the level of production, deviations from the monolinguals' norm were detected in the speech of L1 participants who acquired their L2 in adulthood. In addition, this study revealed that late L2 learners, who acquired their L2 in adulthood, can succeed in acquiring and producing the L2 like natives in the phonetics domain.

In regards to L2 acquisition, the bilinguals in this study achieved native-like speech in some aspects but not in others. For example, A-E and E-A bilinguals both approximated the native-like prosody of *wh*-words in Arabic and English. However, the E-A bilinguals did not achieve native-like speech in Arabic vowels and Arabic voiceless plosives, and A-E bilinguals did not achieve native-like pronunciation in the English /t/ and /k/ sounds.

Regarding L1 attrition, the results from the bilinguals indicated some asymmetric patterns of L1 attrition, with evidence for attrition among the E-A bilinguals (but not among the A-E bilinguals) in the production of *wh*-words and the *VOT* values of voiceless plosives. The production of the shared vowels revealed L1 attrition in both bilingual groups.

When all factors were examined, language use, but not phonetic aptitude (or at least to a much lesser extent), turned out to affect prosody. However, there is no evidence that language use or phonetic aptitude influenced the formants of shared vowels (Study II) as well as the *VOTs* (Study III) of A-E and E-A bilinguals in both L1 attrition and L2 acquisition.

To conclude, L1 phonetic attrition appears to be prevented by L1 use, but only at the suprasegmental level. Some, but not all, of the current segmental results support the SLM and PAM-L2. In addition, there are some areas of pronunciation that may be more vulnerable to attrition than others, and some areas may be more likely to achieve native-likeness than others. The present project revealed that there are some aspects of native-like L2 acquisition that can occur along with L1 attrition, which means that, while they are acquired well in the L2, they are susceptible to attrition in L1. However, this is not the case for all the aspects. Finally, the current study confirms the idea that acquiring a language from birth is not sufficient to ensure native-likeness in the production of bilingual speech.

References

Abrahamsson, N. and Hyltenstam, K. 2009. Age of onset and native likeness in a second language: listener perception versus linguistic scrutiny. *Language Learning*, 59, pp.249-306.

Abu-Rabia, S. & Kehat, S. 2004. The critical period in second language pronunciation: is there such a thing? Ten case studies. *Educational Psychology*, 24 (1), pp.77-98.

Adank, P., Van Hout, R. & Smits, R. 2004. An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical Society of America*, 116 (3), pp.1729-1738.

Ahn, M.J. 2019. Effects of F0 and VOT on the Identification of Korean Word-Initial Oral and Nasal Stops. *언어과학연구*, *90*, pp.203-218.

Al-Amer, A. 2019. *Phonological acquisition and development in Arabic-English bilingual children*. PhD dissertation. Birkbeck: University of London.

Alanazi, S. 2018. *The Acquisition of English stops by Saudi L2 Learners*. Doctoral dissertation. Colchester, Essex: University of Essex.

Al-Ani, S.H. 1970. *Arabic phonology*; *an Acoustical and Physiological Investigation*. The Hague, Mouton: Walter de Gruyter.

AlDahri, S.S. 2013. A study for the effect of the emphaticness and language and dialect for voice onset time (VOT) in Modern Standard Arabic (MSA).

Alghamdi, M. 1990. *Analysis, synthesis and perception of voicing in Arabic*. PhD thesis. Department of Linguistic Science, University of Reading.

Alghamdi, M. 2006. Voice print: voice onset time as a model. *Arab Journal for Security Studies and Training*, 21 (42), pp.89-188.

Alghamdi, M. 1998. A spectrographic analysis of Arabic vowels: a cross-dialect study. *Journal of King Saud University*, 10 (1), pp.3-24.

Almbark, R. 2012. *The perception and production of SSBE vowels by Syrian Arabic learners: the foreign language model.* PhD dissertation, York: University of York.

Alosh, M. 2005. *Using Arabic: a guide to contemporary usage*. Cambridge: Cambridge University Press.

Alotaibi, M. 2013. A problem with wh-questions in Modern Standard Arabic. *Proceedings of Language at the University of Essex*. pp.1-8.

Alotaibi, Y. & Husain, A. 2009. Formant based analysis of spoken Arabic vowels. In: Young-hoon, L, Tai-hoon, K, Wai-chi, F & Dominik, Ś (eds.). *European Workshop on Biometrics and Identity Management.* Berlin, Heidelberg: Springer. pp.162-169.

Ammerlaan, T. 1996. "*You get a bit wobbly...": exploring bilingual lexical retrieval processes in the context of first language attrition*. Nijmegen: Ph.D. thesis, Katholieke Universiteit Nijmegen.

Aoun, J. & Benmamoun, E. 1998. Minimality, reconstruction, and PF movement. *Linguistic Inquiry*, 29 (4), pp.569-597.

Aoyama, K., Flege, J.E., Guion, S.G., Akahane-Yamada, R. & Yamada, T. 2004. Perceived phonetic dissimilarity and L2 speech learning: /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, pp.233-250.

Arvaniti, A., Ladd, D.R. & Mennen, I. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics*, 26, pp.3-25.

Baayen, R.H. 2008. *Analyzing linguistic data: a practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Baker, C. 2011. *Foundations of bilingual education and bilingualism*. vol. 79. Bristol: Multilingual Matters.

Barr, D.J., Levy, R., Scheepers, C. & Tily, H.J. 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*. [E-journal]. 68, pp.255-278. Available at: http://dx.doi.org/10.1016/j.jml.2012.11.001. Accessed 10th January 2016.

Bartels, C. 2014. *The intonation of English statements and questions: a compositional interpretation*. England: Routledge. Taylor & Francis Group.

Bates, D., Maechler, M., Bolker, B. & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. [E-journal]. 67 (1), pp. 1–48. Available at: http://dx.doi.org/10.18637/jss.v067.i01. Accessed 24th February 2017.

Belkaid, Y. 1984. Arabic vowels, modern literature, spectrographic analysis. *Phonetic Works Strasbourg Institution*, *16*, pp.217-240.

Bergmann, C., Sprenger, S.A. & Schmid, M.S. 2015. The impact of language co-activation on L1 and L2 speech fluency. *Acta Psychologica*, 161, pp.25-35.

Bergmann, C., Nota, A., Sprenger, S.A. & Schmid, M.S. 2016. L2 immersion causes nonnative-like L1 pronunciation in German attriters. *Journal of Phonetics*, 58, pp.71-86.

Best, C.T. 1995. A direct realist view of cross-language speech perception. In: Strange, W. (ed.). *Speech perception and linguistic experience: issues in cross-language research*. Timonium, MD: York Press. pp.171-232.

Best, C.T. & Tyler, M.D. 2007. Nonnative and second-language speech perception: commonalities and complementarities. In Bohn, O-S. & Munro, M.J. (eds.). *Language experience in second language learning: in honour of James Emil Flege.* Amsterdam: John Benjamins. pp.13-34.

Best, C.T., McRoberts, G.W. & Sithole, N.M. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14 (3), p.345.

Bialystok, E. 2001. *Bilingualism in development: language, literacy, & cognition*. New York: Cambridge University Press.

Boersma, P. & Weenink, D. 2016. Praat: doing phonetics by computer. [Computer program]. Version 6.0.28. Accessed 10th January 2016.

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), pp.131-158.

Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19 (04), pp.447-465.

Bongaerts, T. 1999. Ultimate attainment in L2 pronunciation: the case of very advanced late L2 learners. In: Birdsong, D. (ed.). *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Lawrence Erlbaum. pp.133-160.

Bongaerts, T., Mennen, S. & Van der Slik, F. 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54, pp.298-308.

Bongaerts, T., Planken, B. & Schils, E. 1995. Can late starters attain a native accent in a foreign language: A test of the critical period hypothesis. In: Singleton, D. & Lengyel, Z. (eds). *The age factor in second language acquisition*. Clevedon: Multilingual Matters. pp.30-50.

Bonn, O.-S. & Flege, J.E. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, pp.303-328.

Bowen, J.D. 1972. Contextualizing pronunciation practice in the ESOL classroom. *TESOL Quarterly*,6, pp.83-94.

Brennan, E.M., Ryan, E.B. & Dawson, W.E. 1975. Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research*, 4 (1), pp.27-36.

Bylund, E. 2009. Maturational constraints and first language attrition. Language Learning. [E-journal]. 59 (3), pp.687-715. Available at: http://dx.doi.org/10.1111/j.1467-9922.2009.00521.x.

Bylund, E., Abrahamsson, N. & Hyltenstam, K. 2009. The role of language aptitude in first language attrition: the case of pre-pubescent attriters. *Applied Linguistics*, 31 (3), pp.443-464.

Cabrelli, J., Luque, A. and Finestrat-Martínez, I., 2019. Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception. *Journal of Phonetics*, *73*, pp.55-69.

Garamazza, A., Yeni-Komshian, G. H., Zurif, E.B. & Carbone, E. 1973. The acquisition of a new phonological contrast: the case of stop consonants in French–English bilinguals. *Journal of the Acoustical Society of America*, 54, pp.421-428.

Carroll, J.B. & Sapon, S.M. 2002. *Modern Language Aptitude Test: manual 2002 edition*. [Online]. Bethesda, MD: Second Language Testing, Inc. Available at: http://lltf.net/wp-content/uploads/2011/09/Test_mlat.pdf. Accessed 2nd October 2015.

Carroll, J.B. 1993. *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Cebrian, J. 2006. Experience and the use of duration in the categorization of L2 vowels. *Journal of Phonetics*, 34, pp.372-387.

Celce-Murcia, M., Brinton, D.M. and Goodwin, J.M., 1996. *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

Celce-Murcia, M., Brinton, D.M. & Goodwin, J.M. 2010. *Teaching pronunciation hardback with audio CDs (2): a course book and reference guide*. Cambridge University Press.

Chang, C.B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40 (2), pp.249-268.

Chang, C.B. 2013. A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, 41, pp.520-533.

Chen, R.T.H., Bennett, S. & Maton, K. 2008. The adaptation of Chinese international students to online flexible learning: two case studies. *Distance Education*, 29 (3), pp.307-323.

Chen, S.X., Benet-Martinez, V. & Bond M. 2008. Bicultural identity, bilingualism, and psychological adjustment in multicultural societies: immigration-based and globalization-based acculturation. *Journal of Personality*, 76 (4), pp.803-838.

Cho, T., Whalen, D.H. & Docherty, G. 2019. Voice onset time and beyond: exploring laryngeal contrast in 19 languages. *Journal of Phonetics*, 72, pp.52-65.

Choi, J., Broersma, M. & Cutler, A. 2017. Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences*, 114 (28), pp.7307-7312.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. 1975. *The logical structure of linguistic theory*. Cambridge, MA: MIT Press.

Chomsky, N. 1981. *Lectures on Government and Binding:* The Pisa lectures. Dordrecht: Foris, Berlin Mounton.

Chomsky, N. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. & Halle, M. 1968. *The sound pattern of English*. New York; Evanston; London: Harper & Row.

Clopper, C. G. (2009). Computational methods for normalizing acoustic vowel data for talker differences. *Language and Linguistics Compass*, *3*(6), 1430-1442.

Clyne, M. 1997. Some of the things trilinguals do. *International Journal of Bilingualism*, 1 (2), pp.95-116.

Clyne, M. (ed). 2012. *Pluricentric languages: differing norms in different nations.* vol. 62. Walter de Gruyter.

Cohen, A. 1975. Forgetting a second language. *Language Learning*, 25, pp.127-138.

Cook, V. 2007. The goals of ELT. In: *International handbook of English language teaching*. Boston, MA: Springer. pp.237-248.

Cook, V. 2003. The changing L1 in the L2 user's mind. In: Cook, V. (ed.). *Effects of the second language on the first*. Clevedon: Multilingual Matters.

Davidson, L. 2006. Phonology, phonetics, or frequency: influences on the production of non-native sequences. *Journal of Phonetics*, 34, pp.104-137.

Davis, C. 1999. Will the use of videos designed for the purpose of teaching English pronunciation improve the learners' production of discrete sounds by at least 80% over a 12-week period? In: *Action Research Monographs*. Pennsylvania Action Research Network. Section 353 Project of the Pennsylvania Department of Education, Bureau of Adult Basic and Literacy Education. ERIC Document [ED440 219].

De Bot, K. & Stoessel, S. 2000. In search of yesterday's words: reactivating a long-forgotten language. *Applied Linguistics*, 21 (3), pp.333-353.

De Clercq, B. & Housen, A. 2019. The development of morphological complexity: a cross-linguistic study of L2 French and English. *Second Language Research*, 35 (1), pp.71-97.

De Jong, K. 2004. Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics*, 32, pp.493-516.

De Bot, K. & Mailfert, K. 1982. The teaching of intonation: fundamental research and classroom applications. *TESOL Quarterly*, 16, pp.71-77.

De Jong, K. 1994. Initial tones and prominence in Seoul Korean. In: Lee, S.-H. & Jun, S.-A. (eds.). *Papers from the Linguistics Laboratory Volume 43.* Columbus, OH: Ohio State University, pp.1-14.

De Jong, K. & Zawaydeh, B. 2002. Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30, pp.53-75.

De Leeuw, E. 2009. *When your native language sounds foreign: a phonetic investigation into first language attrition*. PhD thesis. Edinburgh, UK: Queen Margaret University.

De Leeuw, E., Mennen, I. & Scobbie, J.M. 2011. Singing a different tune in your native language: first language attrition of prosody. *International Journal of Bilingualism*, 16 (1), pp.101-116.

De Leeuw, E., Schmid, M.S. & Mennen, I. 2010. The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, 13 (1), pp.33-40.

De Leeuw, E.D., Hox, J. & Dillman, D. 2012. *International handbook of survey methodology*. New York: Routledge.

De Leeuw, E., Mennen, I. & Scobbie, J.M. 2013. Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism*, 17 (6), pp.683-700.

Defense Language Institute. 1974. *A contrastive study of English and Arabic*. [Online]. Monterey, CA: Available at: https://files.eric.ed.gov/fulltext/ED105761.pdf. Accessed 26th March 2016.

Denes, P. 1955. Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America*, 27 (4), pp.761-764.

Derwing, T.M., Munro, M.J. & Wiebe, G.E. 1998. Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, pp.393-410.

Deterding, D. 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, 27 (1-2), pp.47-55.

Deterding, D. 2006. The pronunciation of English by speakers from China. *English World-Wide*, 27 (2), pp.175-198.

Deuchar, M. & Clark, A. 1996. Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics*, 24, pp.351-365.

Diamond, M.C., Scheibel, A.B., Murphy, G.M. & Harvey, T. 1985. On the brain of a scientist: Albert Einstein. *Experimental Neurology*, 88 (1), pp.198-204.

Docherty, G. 1992. *The timing of voicing in British English obstruents*. Berlin & New York: Foris Publications.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In: Doughty, C.J. & Long, M.H. (eds.). *The handbook of second language acquisition*. Oxford, UK: Blackwell. pp.589-630.

Dusseldorp, E. & Schmid, M. 2010. Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. Groningen: Rijksuniversiteit/TNO, Quality of Life & Leiden University.

Ecke, P. 2004. Language attrition and theories of forgetting: a cross-disciplinary review. *International Journal of Bilingualism*, 8 (3), pp.321-354.

Eckman, F.R. 1981. On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition*, 4, pp.18-30.

Eckman, F.R. 1987. Markedness and the contrastive analysis hypothesis. *Language Learning*, 27 (2), pp.315-330.

El Zarka, D. 2017. Arabic intonation. *Oxford handbooks online.* [Online]. Available at: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-77. Accessed 16th February 2018.

El Zarka, D. 1997. *Prosodische phonologie des Arabischen*. PhD dissertation. Karl-Franzens-Universität Graz.

Els, T. & Bot, K. 1987. The role of intonation in foreign accent. *The Modern Language Journal*, 71 (2), pp.147-155.

Escudero, P. 2005. *Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization*. PhD dissertation. the Netherlands: Utrecht University.

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

Escudero, P. & Boersma, P. 2002. The subset problem in L2 perceptual development: multiple-category assimilation by Dutch learners of Spanish. In: Skarabela, B., Fish, S. & Do, A.H.J. (eds). *Proceedings of the 26th annual Boston University conference on language development*. Somerville: Cascadilla Press. pp.208-219.

Escudero, P. & Boersma, P. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26, pp.551-585.

Fabbro, F. 2001. The bilingual brain: bilingual aphasia. *Brain and Language*, 79, pp.1842-1847.

Fabiano-Smith, L. & Barlow, J.A. 2010. Interaction in bilingual phonological acquisition: evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism*, 13 (1), pp.81-97.

Fabiano-Smith, L. & Goldstein, B.A. 2010. Phonological acquisition in bilingual Spanish–English speaking children. *Journal of Speech, Language, and Hearing Research*, 53 (1), pp.160-178.

Fakih, A.H.A. 2015. Wh-questions in Hodeidi Arabic: a phase-based approach. *Theory and Practice in Language Studies*, 5 (4), p.773.

Favier, S., Wright, A., Meyer, A. & Huettig, F. 2019. Proficiency modulates between – but not within – language structural priming. *Journal of Cultural Cognitive Science*, 3 (1), pp.105–124.

Flege, J.E. 1980. Phonetic approximation in second language acquisition. *Language Learning*, 30 (1), pp.117-134.

Flege, J.E. 1987a. The production of "new" and "similar" phonemes in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, pp.47-65.

Flege, J.E. 1987b. A critical period for learning to pronounce foreign languages? *Applied Linguistics*, 8, pp.162-177.

Flege, J.E. 1988. The production and perception of foreign language speech sounds. In: *Human communication and its disorders–a review*. pp.224-401.

Flege, J.E. 1989. The instrumental study of L2 speech production: some methodological considerations. *Language Learning*, 37 (2), pp.285-296.

Flege, J.E. 1991a. Perception and production: the relevance of phonetic input to L2 phonological learning. In: Huebner, T. & Ferguson, C.A. (eds.). *Cross currents in second* *language acquisition and linguistic theory*. vol 2. Amsterdam: John Benjamins Publishing. pp.249-289.

Flege, J.E. 1991b. The intelligibility of English vowels spoken by British and Dutch talkers. In: Kent, R. (ed.). *Intelligibility in speech disorders.* Amsterdam: John Benjamins. pp.157-231.

Flege, J.E. 1993. Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America*, 93 (3), pp.1589-1608.

Flege, J.E. 1995. Second language speech learning, theory, findings, and problems. In: Strange, W. (ed.). *Speech perception and linguistic experience: issues in cross-language research*. Timonium, MD: York Press. pp.229-273.

Flege, J.E. 1999. Age of learning and second-language speech. In: Birdsong, D. (ed.). *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Erlbaum. pp.101-132.

Flege, J.E. 1980. Phonetic approximation in second language acquisition. *Language Learning* 30 (1), pp.117-134.

Flege, J.E. 1989. The instrumental study of L2 speech production: some methodological considerations. *Language Learning*, 37 (2), pp.285-296.

Flege, J.E. 1992. Speech learning in a second language. In: Ferguson, C.A., Menn, L. & Stoel-Gammon, C. (eds.). *Phonological development: Models, research, implications*. Timonium, MD: York Press. pp.565-604.

Flege, J.E. 1992. The intelligibility of English vowels spoken by British and Dutch talkers. *Intelligibility in Speech Disorders: Theory, Measurement, and Management*, 1, pp.157-232.

Flege, J.E. 1995. Second language speech learning, theory, findings, and problems. In: Strange, W. (ed.). *Speech perception and linguistic experience: issues in cross-language research*. Timonium, MD: York Press. pp.229-273.

Flege, J.E. 1999. Age of learning and second-language speech. In: Birdsong, D. (ed). *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Erlbaum. pp.101-132.

Flege, J.E. 2003. Assessing constraints on second-language segmental production and perception. In: Meyer, A. & Schiller, N. (eds). *Phonetics and phonology in language comprehension and production*. Berlin: Mouton de Gruyter. pp.319-355.

Flege, J.E. 2009. Give input a chance! In: Piske, T., Young-Scholten, M. (eds.). *Input matters in SLA*. Clevedon: Multilingual Matters. pp.175-190.

Flege, J.E., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Tsukada, K. 2006. Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, 34, pp.153-175.

Flege, J.E. & Eefting, W. 1987a. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, pp.185-202.

Flege, J.E. & Eefting, W. 1987b. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, pp.67-83.

Flege, J.E. & Fletcher, K.L. 1992. Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America*, 91 (1), pp.370-389.

Flege, J.E. & Hillenbrand, J. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America*, 76 (3), pp.708-721.

Flege, J.E. & MacKay, I.R.A. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, pp.1-34.

Flege, J.E., MacKay, I.R. & Meador, D. 1999. Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106 (5), pp.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), pp.3125-3134.

Flege, J.E. & Port, R. 1981. Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24, pp.125-146.

Francis, A.L., Giocca, V., Ma, L. & Fenn, K. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36, pp.268-294.

Fraser, H. 2010. Cognitive theory as a tool for teaching second language pronunciation. *Fostering Language Teaching Efficiency through Cognitive Linguistics*, 6, pp.357-379.

Gardner, H. 1983/1994. *Frames of mind: the theory of multiple intelligences*. New York: Basic Books.

Gass, S. 1983. Language transfer and universal grammatical relations. In: Gass, S. & Selinker, L. (eds.). *Language transfer in language learning*. Rowley, MA: Newbury House. pp.69-82.

Gass, S. 1996. Second language acquisition and the role of language transfer. In: Ritchie, W. & Bhatia, T. (eds.). *Handbook of second language acquisition.* San Diego: Academic Press. pp.317-345.

Gass, S. & Selinker, L. 2008. *Second language acquisition: an introductory course.* London: Routledge.

Gathercole, S.E., Willis, C.S., Baddeley, A.D. & Emslie, H. 1994. The children's test of nonword repetition: a test of phonological working memory. *Memory*, 2 (2), pp.103-127.

Geschwind, N. & Galaburda, A.M. 1985. Cerebral lateralization: biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. *Archives of Neurology*, 42 (5), pp.428-459.

Ghazali, S. 1979. Du statut des voyelles en arabe. *Analyses & Théorie*, 2/3, pp.199-219.

Gnevsheva, K. 2017. Within-speaker variation in passing for a native speaker. *International Journal of Bilingualism*, 21 (2), pp.213-227.

Goad, H., White, L. & Steele, J. 2003. Missing inflection in L2 acquisition: defective syntax or LI-constrained prosodic representations? *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 48 (3-4), pp.243-263.

Goetz, P.J. 2003. The effects of bilingualism on theory of mind development. *Bilingualism, Language and Cognition*, 6, pp.1-15.

Goral, M. 2004. First-language decline in healthy aging: implications for attrition in bilingualism. *Journal of Neurolinguistics*, 17 (1), pp.31-52.

Gordeeva, O. 2006. *Language interaction in the bilingual acquisition of sound structure: a longitudinal study of vowel quality, duration and vocal effort in pre-school children speaking Scottish English and Russian.* PhD dissertation. Edinburgh: Queen Margaret University.

Gordon, E. 1989. *Advanced measures of music audiation.* Chicago: GIA Publications.

Grosjean, F. 1982. *Life with two languages: an introduction to bilingualism.* Cambridge, MA: Harvard College.

Grosjean, F. 1989. Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, pp.3-15.

Grosjean, F. 2001. The bilingual's language modes. In: Nicol, J. (ed.). *One mind, two languages: bilingual language processing.* Oxford: Blackwell. pp.1-22.

Guion, S.G. 2003. The vowel systems of Quichua–Spanish bilinguals. *Phonetica*, 60 (2), pp.98-128.

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 (5), pp.2711-2724.

Guion, S.G., Flege, J.E., Liu, S.H. & Yeni-Komshian, G.H. 2000. Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, 21 (2), pp.205-228.

Guiora, A.Z. 1990. A psychological theory of second language pronunciation. *Toegepaste Taalwetenschap in Artikelen*, 37 (1), pp.15-23.

Hamers, J.F., Blanc, M. & Blanc, M.H. 2000. *Bilinguality and bilingualism*. Cambridge University Press.

Hawkins, S. & Midgley, J. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, 35 (2), pp.183-199.

Hedberg, N. & Sosa, J.M. 2002. The prosody of questions in natural discourse. *Proceedings of Speech Prosody 2002: an international conference.* Aix-en-Provence, France, 11-13 April 2002. pp.375-378.

Hedberg, N., Sosa, J.M., Görgülü, E. & Mameni, M., 2010. Prosody and pragmatics of wh-interrogatives. In: *2010 Annual Meeting of the Canadian Linguistics Society, Proceedings*.

Hellmuth, S. & Almbark, R. 2015. The prosody of interrogatives: evidence from spoken Arabic dialects. *ProVar: Workshop on Prosodic Variation.* Universidade de Lisboa.

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), pp.3099-3111.

Hinkel, E. 2006. Current perspectives on teaching the four skills. *TESOL Quarterly*, 40, pp.109-131.

Hinton, M. 2012. *The identification and testing of cognitive and affective factors in language learning aptitude with Polish students of English*. [Unpublished manuscript]. PhD thesis. Łódź: University of Łódź.

Hinton, M. 2013. An aptitude for speech: the importance of mimicry ability in foreign language pronunciation. In: *Teaching and researching English accents in native and non-native speakers.* Berlin: Springer. pp.103-111.

Hinofotis, F. & Baily, K. 1980. American undergraduate reaction to the communication skills of foreign teaching assistants. *TESOL* "*80: Building Bridges: Research and Practice in TESL*". Alexandria, VA.

Hopp, H. & Schmid, M.S. 2014. Perceived foreign accent in first language attrition and second language acquisition: the impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, 34 (2), pp.361-394.

Hopp, H. 2010. Ultimate attainment in L2 inflection: performance similarities between non-native and native speakers. *Lingua*, 120 (4), pp. 901-931.

Horgues, C. 2010. *Prosodie de l'accent français en anglais et perception par des auditeurs anglophones*. [Unpublished manuscript]. PhD dissertation. University Denis-Diderot-Paris.

Houston, D. 2011. Infant speech perception. In: Seewald, R. & Tharpe, A.M. (eds.). *Comprehensive handbook of pediatric audiology.* San Diego, CA: Plural Publishing. pp.47-62.

Hsui, V.Y. 1996. Bilingual but not biliterate: case of a multilingual Asian society. *Journal of Adolescent & Adult Literacy*, 39 (5), pp.410-414.

Hume, E. 2004. Deconstructing markedness: a predictability-based approach. *Proceedings of the Berkeley Linguistic Society*, 30, pp.182-198.

Ingram, J.C. & Park, S.G. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25 (3), pp.343-370.

Iverson, P. & Kuhl, P.K. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *The Journal of the Acoustical Society of America*, 97 (1), pp.553-562.

Iverson, P. & Kuhl, P.K., 1996. Influences of phonetic identification and category goodness on American listeners' perception of/r/and/l. *The Journal of the Acoustical Society of America*, 99 (2), pp.1130-1140.

Jesry, M. 1996. *Some cognitively controlled coarticulatory effects in Arabic and English, with particular reference to voice onset time*. University of Essex.

Jia, G., Aaronson, D. & Wu, Y. 2002. Long-term language attainment of bilingual immigrants: predictive variables and language group differences. *Applied Psycholinguistics*, 23 (4), pp.599-621.

Jilka, M. 2007. Different manifestations and perceptions of foreign accent in intonation. In: Trouvain, J. & Gut, U. (eds). *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter. pp.76-96.

Jilka, M. 2009. Assessment of phonetic ability. In: Dogil, G. & Reiterer, S. (eds.). *Language talent and brain activity.* Berlin: Mouton de Gruyter. pp.1-16.

Jilka, M., Anufryk, V., Baumotte, H., Lewandowska, N., Rota, G. & Reiterer, S. 2008 Assessing individual talent in second language production and perception. In: Rauber, A. Watkins, M. & Baptista, B. (eds.). *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech.* Florianopolis: University of Santa Catarina. pp.224-239.

Jilka, M., Lewandowska, N. & Rota, G. 2011. Investigating the concept of talent in phonetic performance. In: Wrembel, M., Kul, K. & Dziubalska-Kolaczyk, K. (eds.). *Achievements and perspectives in SLA of speech: new sound 2010.* Frankfurt am Main: Peter Lang. pp.171-180.

Johnson, J.S. & Newport, E.L. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, pp.60-99.

Jun, S.A. & Cowie, I. 1994. Interference for 'new' versus 'similar' vowels in Korean speakers of English. In: Lee, S.-H. & Jun, S.-A. (eds.). *Papers from the Linguistics Laboratory Volume 43*. Columbus, OH: Ohio State University, pp.117-130.

Keating, P., Linker, W. & Huffman, M. 1983. Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics, 11* (1), pp.277-290.

Keating, P.A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language*, 60 (2), pp.286-319.

Khattab, G. 2000. VOT production in English and Arabic bilingual and monolingual children. *Leeds working papers in linguistics and phonetics*, 8, pp.95-122.

Khattab, G. 2002. /l/ production in English–Arabic bilingual speakers. *International Journal of Bilingualism*, 6, pp.335-353.

Khattab, G. 2002. *Sociolinguistic competence and the bilingual's adoption of phonetic variants: auditory and instrumental data from English-Arabic bilinguals*. Doctoral dissertation. Leeds: University of Leeds.

Khattab, G. 2007. Variation in vowel production by English–Arabic bilinguals. *Laboratory Phonology*, 9, pp.383-410.

Khattab, G. & Al-Tamimi, J. 2008. Durational cues for gemination in Lebanese Arabic. *Language and Linguistics*, 22, pp.39-55.

Khomitsevich, O. 2008. *Dependencies across phases: from sequence of tense to restrictions on movement.* PhD Dissertation. Utrecht: Netherlands Graduate School of Linguistics.

Kim, M. 2004. Correlation between VOT and F0 in the perception of Korean stops and affricates. *Proceedings of the 8th International Conference on Spoken Language*

Processing (INTERSPEECH-2004). Jeju Island, Korea. International Speech Communication Association. pp.49-52.

Kim, S., Kim, J. and Cho, T., 2018. Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics*, *71*, pp.65-80.

Klatt, D.H. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, 18 (4), pp.686-706.

Kopczynski, A. & Meliani, R. 1993. The vowels of Arabic and English. *Papers and Studies in Contrastive Linguistics*, 27, pp.183-192.

Köpke, B. 2004. Neurolinguistic aspects of attrition. *Journal of Neurolinguistics*, 17 (1), pp.3-30.

Köpke, B. 2007. Language attrition at the crossroads of brain, mind, and society. In Köpke, B., Schmid, M.S., Keijzer, M. & Dostert, S. (eds.). *Language attrition: theoretical perspectives.* Amsterdam: John Benjamins. pp.9-38.

Köpke, B. & Genevska-Hanke, D. 2018. First language attrition and dominance: same or different? *Frontiers in Psychology*, 9, p.1963.

Köpke, B. & Schmid, M. 2007. Bilingualism and attrition. Groningen: Rijksuniversiteit/Le Mirail: Laboratoire de Neuropsycholinguistique, Université de Toulouse.

Köpke, B. & Schmid, M.S. 2004. Language attrition: the next phase. In: Schmid, M.S.,

Köpke, B., Keijzer, M. & Weilemar, L. (eds.). *First language attrition: interdisciplinary perspectives on methodological issues*. Amsterdam: John Benjamins. pp.1-43.

Krashen, S.D. 1981. *Second language acquisition and second language learning*. Oxford: Oxford University Press.

Krashen, S.D. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon.

Kuhl, P.K. 1991. Human adults and human infants show a ‘perceptual magnet effect’ for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50 (2), pp.93-107.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N. & Lindblom, B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255 (5044), pp.606-608.

Labov, W. 2006. *The social stratification of English in New York City*. Cambridge: Cambridge University Press.

Ladd, R.D. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Ladefoged, P. & Johnson, K. 2011. *A course in phonetics*. Boston, MA: Wadsforth.

Lado, R. 1957. *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lambert, R.D. & Freed, B.F. 1982. *The loss of language skills.* Rowley, MA: Newbury House.

Lambrecht, K. & Michaelis, L., 1998. Sentence accent in information questions: default and projection. *Linguistics and Philosophy*, 21, pp.477-544.

Language Learning and Testing Foundation. 2019. *Modern Language Aptitude Test.* [Online] Available at: http://lltf.net/aptitude-tests/language-aptitude-tests/modern-language-aptitude-test-2/. Accessed 10th February 2015.

Language Learning and Testing Foundation. 2019. *Pimsleur Language Aptitude Battery*. [Online]. Available at: http://lltf.net/aptitude-tests/language-aptitude-tests/pimsleur-language-aptitude-battery/. Accessed 10th February 2015.

Larsen-Freeman, D. & Long, M. 1991. *An introduction to second language acquisition research.* London: Longman.

Larson, G. 1984. The Far Side gallery 1. London: Warner Books.

Larson, G. 1992. *The far side observer*. Warner Books.

Leather, J. 1987. F0 pattern inference in the perceptual acquisition of second language tone in sound patterns in second language acquisition. *Studies on Language Acquisition (SOLA)*, 5, pp.59-80.

Leal, T. & Slabakova, R. 2019. The relationship between L2 instruction, exposure, and the L2 acquisition of a syntax–discourse property in L2 Spanish. *Language Teaching Research*, 23 (2), pp.237-258.

Lee, S.A.S. & Iverson, G.K. 2012. Stop consonant productions of Korean–English bilingual children. *Bilingualism: Language and Cognition*, 15 (2), pp.275-287.

Lee, T. 2011. Grammatical knowledge of Korean heritage speakers: early vs. late bilinguals. *Linguistic Approaches to Bilingualism*, 1 (2), pp.149-174.

Li, Y. 2016. The role of Mandarin speakers' mimicry ability in their accurate pronunciation of French. *The Asian EFL Journal*, 95, pp.63-74.

Lightbown, P.M. & Spada, N. 2006. *How languages are learned: an introduction to the main theories of first and second language acquisition.* Oxford: Oxford University Press.

Lisker, L. & Abrahamson, A.S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, pp.384-422.

Lisker, L. & Abrahamson, A.S. 1967. Some effects of context on voice onset time in English stops. *Language and Speech*, 10 (1), pp.1-28.

Long, M.H. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, pp.251-285.

Lord, G. 2005. (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88, pp.257-267.

MacKay, I.R., Flege, J.E., Piske, T. & Schirru, C. 2001. Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America*, 110 (1), pp.516-528.

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, pp.157-183.

Magen, H.S. 1998. The perception of foreign-accented speech. *Journal of Phonetics*, 26 (4), pp. 381-400.

Major, R.C. 1987a. A model for interlanguage phonology. *Interlanguage Phonology: The Acquisition of a Second Language Sound System*,15 pp.101-124.

Major, R.C. 1987b. English voiceless stop production by Brazilian speakers of English. *Phonetics*, 15, pp.197-222.

Major, R.C. 1992. Losing English as a first language. *The Modern Language Journal*, 76, pp.190-208.

Major, R.C. 2001. *Foreign accent: the ontogeny and phylogeny of second language phonology*. England: Routledge.

Major, R.C. 2010. First language attrition in foreign accent perception. *International Journal of Bilingualism*, 14 (2), pp.163-183.

Mattock, K., Molnar, M., Polka, L. & Burnham, D. 2008. The developmental course of lexical tone perception in the first year of life. *Cognition*, 106 (3), pp.1367-1381.

Mayr, R., Price, S. & Mennen, I., 2012. First language attrition in the speech of Dutch–English bilinguals: the case of monozygotic twin sisters. *Bilingualism: Language and Cognition*, 15 (4), pp.687-700.

McGory, J.T. 1997. *Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers.* PhD dissertation. Columbus: Ohio State University.

Meisel, J. 2011. *First and second language acquisition: parallels and differences.* Cambridge: Cambridge University Press.

Mennen, I. 1999. *Second language acquisition of intonation: the case of Dutch near-native speakers of Greek*. PhD dissertation. Scotland: University of Edinburgh.

Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32 (4), pp.543-563.

Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. *Trends in Linguistics Studies and Monographs*, 186, p.53.

Mirman, D. 2014. *Growth curve analysis and visualization using R. The R series*. Boca Raton: CRC Press.

Mirman, D., Dixon, J.A. & Magnuson, J.S. 2008. Statistical and computational models of the visual world paradigm: growth curves and individual differences. *Journal of Memory and Language*, 59 (4), pp.475-494.

Mitchell, T.F. 1990. *Pronouncing Arabic*. 2 vol. Oxford University Press.

Mitleb, F.M. 2009. Voice onset time of Jordanian Arabic stops. *3rd International Conference on Arabic Language Processing (CITALA'09)*. 4-5 May, 2009. Rabat, Morocco.

Mitleb, F. 1982. Voicing effect on vowel duration is not an absolute universal. *The Journal of the Acoustical Society of America*, 71 (S1), pp.S23-S23.

Mitleb, F. 1984. Vowel length contrast in Arabic and English–a spectrographic test. *Journal of Phonetics*, 12 (3), pp.229-235.

Morrison, G.S. 2009. L1 Spanish speakers' acquisition of the English /i/-/I/ contrast II: perception of vowel inherent spectral change. *Language and Speech*, 52, pp.437-462.

Moyer, A. 1999. Ultimate attainment in L2 phonology: the critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21 (1), pp.81-108.

Moyer, A. 2004. *Age, accent and experience in second language acquisition*. Clevedon: Multilingual Matters.

Moyer, A. 2007. Do language attitudes determine accent? A study of bilinguals in the USA. *Journal of Multilingual and Multicultural Development*, 28 (6), pp.502-518.

Moyer, A. 2009. Input as a critical means to an end: quantity and quality of experience in L2 phonological attainment. *Input matters in SLA*, pp.159-174.

Müller, N. & Hulk, A. 2001. Crosslinguistic influence in bilingual language acquisition in Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4, pp.1-21.

Munro, M.J. 1993. Productions of English vowels by native speakers of Arabic: acoustic measurements and accentedness ratings. *Language and Speech*, *36* (1), pp.39-66.

Munro, 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, pp.313-334.

Munro, M.J. & Derwing, T.M. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45 (1), pp.73-97.

Murtagh, L. 2003. *Retention and attrition of Irish as a second language: a longitudinal study of general and communicative proficiency in Irish among second level school leavers and the influence of instructional background, language use and attitude/motivation variables*. Dissertation. Netherlands: Rijksuniversiteit Groningen.

Nauck, B. 2001. Intercultural contact and intergenerational transmission in immigrant families. *Journal of Cross-Cultural Psychology*, 32 (2), pp.159-173.

Neufeld, G.G. 1979. Towards a theory of language learning ability. *Language Learning*, 29 (2), pp.227-241.

Newman, D. & Verhoeven, J. 2002. Frequency analysis of Arabic vowels in connected speech. *Antwerp Papers in Linguistics*, 100, pp.77-86.

Novoa, B., Luque, A., Castro, D., Borrego, J.J. & Figueras, A. 1998. Characterization and infectivity of four bacterial strains isolated from brown ring disease-affected clams. *Journal of Invertebrate Pathology*, 71 (1), pp.34-41.

Novoa, L., Fein, D. & Obler, L.K. 1988. Talent in foreign languages: a case study. In: Obler, L. & Fein, D. (eds.). *The exceptional brain: neuropsychology of talent and special abilities*. New York: Guilford. pp.294-302.

Nunan, D. 1999. *Second language teaching and learning.* Boston: Heinle & Heinle.

Odlin, T. 1990. Word-order transfer, metalinguistic awareness and constraints on foreign language learning. In: Van Patten, B. & Lee, J. (eds.). *Second language acquisition – foreign language learning.* Clevedon: Multilingual Matters. pp.95-117.

Olshtain, E. 1986. The attrition of English as a second language. In: Weltens, B., de Bot, K. & Van Els, T. (eds.). *Language attrition in progress*. Dordrecht: Foris. pp.185-202.

Olshtain, E. 1989. Is second language attrition the reversal of second language acquisition? *Studies in Second Language Acquisition*, 11 (2), pp.151-165.

Oyama, S. 1976. A sensitive period for the acquisition of the non-native phonological system. *Journal of Psycholinguistic Research*, 5, pp.261-283.

Pallier, C. 2007. Critical periods in language acquisition and language attrition. In: Köpke, B., Schmid, M.S., Keijzer, M. & Dostert, S. (eds.). *Language attrition: theoretical perspectives.* Amsterdam: John Benjamins. pp.155-168.

Pan, B. & Berko-Gleason, J. 1986. The study of language loss: models and hypotheses for an emerging discipline. *Applied Psycholinguistics*, 7, pp.193-206.

Paradis, J., Genesee, F. & Brago, M.B. 2011. *Dual language development and disorders: a handbook on bilingualism and second language learning*. 2nd ed. Baltimore, MD: Paul H. Brookes Publishing.

Patkowski, M.S. 1990. Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics*, 11, pp.73-89.

Paratis, M. 2007. L1 attrition features predicted by a neurolinguistic theory of bilingualism. In: Köpke, B., Schmid, M.S., Keijzer, M. & Dostert, S. (eds.). *Language attrition: theoretical perspectives.* Amsterdam: John Benjamins. pp.121-134.

Pavlenko, A. 2003. 'Language of the enemy': foreign language education and national identity. *International Journal of Bilingual Education and Bilingualism*, 6 (5), pp.149-191.

Pavlenko, A. 2005. *Emotions and multilingualism.* Cambridge: Cambridge University Press.

Pavlenko, A. 2008. Russian in post-Soviet countries. *Russian Linguistics*, 32, pp.59-80.

Penfield, W.R. & Roberts, L.L. 1959. *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

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), pp.175-184.

Pierrehumbert, J. & Hirschberg, J.B. 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P.R. Morgan, J. & Pollack, M.E. (eds.). *Intentions in Communication*. Cambridge, MA: MIT Press. pp.271-311.

Piller, I. 2002. Passing for a native speaker: identity and success in second language learning. *Journal of Sociolinguistics*, 6, pp.179-206.

Pimsleur, P. 1968. Aptitude testing. *Language Learning*, 18 (s3), pp.73-77.

Pintrich, P. R. & Schunk, D. H. 2002. *Motivationineducation: Theory, research, and application*. (2nd, ed.). New Jersey: Merrill Prentice Hall.

Piske, T., Flege, J.E., MacKay, I.R. & Meador, D. 2002. The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, 59 (1), pp.49-71.

Piske, T., MacKay, I.R.A. & Flege, J.E. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, pp.191-215.

Piske, T. 2007. Implications of James E. Flege's research for the foreign language classroom. In: Bohon, O.-S. & Munro, M.J. (eds.). *Language experience in second language speech learning. In honor of James Emil Flege*. Amsterdam: John Benjamins. pp.301-314.

Port, R.F. & Mitleb, F.M. 1983. Segmental features and implementation in acquisition of English by Arabic speakers. *Journal of Phonetics*, 11 (3), pp.219-229.

Port, R.F. & Rotunno, R. 1979. Relation between voice-onset time and vowel duration. *The Journal of the Acoustical Society of America*, 66 (3), pp.654-662.

Prochazka, T. 1988. The spoken Arabic of Abū Thōr in al-Ḥasa. *Zeitschrift für arabische Linguistik*, 18, pp.59-76.

Purcell, E.T. & Suter, R.W. 1980. Predictors of pronunciation accuracy: a reexamination. *Language Learning*, 30 (2), pp.271-287.

Radwan, M. 1996. *An experimental investigation of the acoustical temporal correlates of voicing contrast in stop consonants (with reference to Arabic)*. University of Essex.

Rasier, L. & Hiligsmann, P. 2007. Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28, pp.41-66.

Rice, K. 2000. Featural markedness in phonology: variation. *Glot*, 4 (7-8), pp.3-6, 3-7.

Riney, T.J. & Takagi, N. 1999. Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, 49 (2), pp.275-302.

Riney, T.J., Takagi, N., Ota, K. & Uchida, Y. 2007. The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*, 35 (3), pp.439-443.

Rochet, B.L. & Yanmei, F. 1991. Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception. *Canadian Acoustics*, 19 (4), pp.105-106.

RStudio. 2018. *RStudio: Integrated development environment for R*. RStudio. [Online]. Available at: http://www.rstudio.org/. Accessed 24th February 2017.

Sancier, M.L. & Fowler, C.A. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25 (4), pp.421-436.

Schmid, M. 2007. The role of L1 use for L1 attrition. In: Köpke, B., Schmid, M.S., Keijzer, M. & Dostert, S. (eds.). *Language attrition: theoretical perspectives.* Amsterdam: John Benjamins. pp.135-154.

Schmid, M. 2008. Defining language attrition. Groningen: Rijksuniversiteit.

Schmid, M.S. 2009. On L1 attrition and the linguistic system. *EUROSLA Yearbook. [E-journal]. 9, pp.212-244.* Available at: http://dx.doi.org/10.1075/eurosla.9.11sch.

Schmid, M. 2011. *Language attrition.* Cambridge: Cambridge University Press.

Schmid, M.S. 2002. *First language attrition, use and maintenance: the case of German Jews in Anglophone countries.* vol. 24. Amsterdam: John Benjamins Publishing.

Schmid, M.S. & Dusseldorp, E. 2010. Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. *Second Language Research*, 26 (1), pp.125-160.

Schmid, M.S. & Hopp, H. 2014. Comparing foreign accent in L1 attrition and L2 acquisition: rating and rater effects. *Language Testing*, 31 (3), pp.367-388.

Schmid, M.S. & Köpke, B. 2007. Bilingualism and attrition. In: Köpke, B., Schmid, M.S., Keijzer, M., Dostert, S. (eds.). *Language attrition: theoretical perspectives*. Amsterdam/Philadelphia: John Benjamins. pp.1-8.

Schmid, M.S., Köpke, B., Keijzer, M. & Weilemar, L. (eds). 2004. *First language attrition: interdisciplinary perspectives on methodological issues*. vol. 28. John Benjamins Publishing.

Schmid, M.S. 2012. Attrition and multilingualism. *The Encyclopedia of Applied Linguistics*.

Schmid, M.S. 2013. First language attrition. *Linguistic Approaches to Bilingualism*, 3 (1), pp.94-115.

Schneiderman, E.I. & Desmarais, C. 1988. The talented language learner: some preliminary findings. *Interlanguage Studies Bulletin (Utrecht)*, 4 (2), pp.91-109.

Schoonmaker-Gates, E. 2015. On voice-onset time as a cue to foreign accent in Spanish: native and nonnative perceptions. *Hispania*, pp.779-791.

Schumann, J.H. 1997. *The neurobiology of affect in language learning*. Oxford: Blackwell.

Scovel, T. 1969. Foreign accent, language acquisition and cerebral dominance. *Language Learning*

Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance 1. *Language Learning*, 19 (3-4), pp.245-253.

Seliger, H.W., Seliger, H., Shohamy, E. & Shohamy, E.G. 1989. *Second language research methods*. Oxford: Oxford University Press.

Selinker, L. & Lamendella, T. 1978. Two perspectives on fossilization in inter-language learning. *Interlanguage Studies Bulletin,* 3, pp.149-191.

Shamin, F. 1996. Learner resistance to innovation in classroom methodology. In: Coleman, H. (ed.). *Society and the language classroom*. Cambridge: Cambridge University Press. pp.105-121.

Shen, X.S. 1990. Ability of learning the prosody of an intonational language by speakers of a tonal language: Chinese speakers learning French prosody. *IRAL-International Review of Applied Linguistics in Language Teaching*, 28 (2), pp.119-134.

Shimizu, K. 1990. *A Cross-language Study of Voicing Contrasts of Stop Consonants in Asian Languages*, Seibido Publishing, Tokyo.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P. & Hirschberg, J. 1992. ToBI: A standard for labeling English prosody. In *Proceedings of the International Conference on Spoken Language Processing* (ICSLP), Banff, II, pp.867-870.

Simon, E. & Leuschner, T. 2010. Laryngeal systems in Dutch, English, and German: a contrastive phonological study on second and third language acquisition. *Journal of Germanic Linguistics*, 22 (4), pp.403-424.

Simon, E. 2009. Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research*, 25 (3), pp.377-408.

Skehan, P. 1991. Individual differences in second language learning. *Studies in Second Language Acquisition*, 13 (2), pp.275-298.

Skehan, P. 2002. Theorising and updating aptitude. *Individual Differences and Instructed Language Learning*, 2, pp.69-94.

Skehan, P. & Ducroquet, L. 1988. *A comparison of first & foreign language learning ability*. Departament of English for Speakers of Other Languages, University of London.

Sorace, A. 2004. Native language attrition and developmental instability at the syntax-discourse interface: data, interpretations and methods. *Bilingualism: Language and Cognition*, 7 (2), pp.143-145.

Stampe, D. 1979. *A dissertation on natural phonology*. Bloomington: Indiana University Linguistics Club.

Steedman, M. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31 (4), pp.649-689.

Stevens, G. 2006. The age-length-onset problem in research on second language acquisition among immigrants. *Language Learning*, 56 (4), pp.671-692.

Stevens, K.N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In: Davis Jr., E.E. & Denes, P.B. (eds.). *Human communication: a unified view*. New York: McGraw-Hill. pp.51-66.

Stoehr, A., Benders, T., van Hell, J.G. & Fikkert, P. 2017. Second language attainment and first language attrition: the case of VOT in immersed Dutch–German late bilinguals. *Second Language Research*, 33 (4), pp.483-518.

Story, B.H. & Bunton, K. 2015. A spectral filtering method for tracking formants in children's speech. *Proceedings of Meetings on Acoustics 169ASA*, 23 (1), 060002.

Stromswold, K. 1999. The cognitive neuroscience of language acquisition. In: Gazzaniga, M.S. (ed.). *The new cognitive neurosciences.* 2nd ed. Cambridge, MA: Massachusetts Institute of Technology. pp.909-932.

Sundara, M., Polka, L. & Genesee, F. 2006. Language-experience facilitates discrimination of /d-/ in monolingual and bilingual acquisition of English. *Cognition*, 100 (2), pp.369-388.

Suter, R.W. 1976. Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26 (2), pp.233-253.

Syrdal, A.K. & Gopal, H.S. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, 79 (4), pp.1086–1100.

Syrdal, A.K. & Jilka, M. 2003. Exploration of question intonation in read American English. *The Journal of the Acoustical Society of America*, 114 (4), p.2397.

Thompson, I. 1991. Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41 (2), pp.177-204.

T'Hart, J. & Collier, R. 1975. Integrating different levels of intonation analysis. *Journal of Phonetics*, *3*, pp.235-255.

Tobin, S.J., Nam, H. & Fowler, C.A., 2017. Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics*, *65*, pp.45-59.

Tomasello, M. 1998. Introduction: a cognitive–functional perspective on language structure. In: Tomasello, M. (ed.). *The new psychology of language: cognitive and functional approaches to language structure.* Mahwah, NJ: Lawrence Erlbaum Associates. pp.vii-xxiii.

Trofimovich, P. & Baker, W. 2006. Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28 (1), pp.1-30.

Tsimpli, I.M. & Papadopoulou, D. 2009. Aspect and the interpretation of motion verbs in L2 Greek. *Representational deficits in SLA. Studies in honor of Roger Hawkins*, pp.187-227.

Tsimpli, I., Sorace, A., Heycock, C. & Filiaci, F. 2004. First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, 8 (3), pp.257-277.

Ueyama, M. & Jun, S.A. 1996. Focus realization of Japanese English and Korean English intonation. *UCLA Working Papers in Phonetics*, pp.110-125.

Ueyama, M. 1997. The phonology and phonetics of second language intonation: the case of 'Japanese English'. *Fifth European Conference on Speech Communication and Technology,* 5, pp.2411-2414. .

Ulbrich, C. & Ordin, M. 2014. Can L2-English influence L1-German? The case of post-vocalic /r/'. *Journal of Phonetics*, 45, pp.26-42.

United Nations 2017. *The International Migration Report 2017 (Highlights).* Available at: https://www.un.org/development/desa/publications/international-migration-report-2017.html. Accessed 18th March 2019.

Van Els, T. 1986. An overview of European research on language attrition. In: Weltens, B., de Bot, K. & Van Els, T. (eds.). *Language attrition in progress.* Dordrecht: Foris. pp.3-18.

Ventureyra, V.A., Pallier, C. & Yoo, H.Y. 2004. The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics*, 17 (1), pp.79-91.

Verdugo, D.R. & Trillo, J.R. 2005. The pragmatic function of intonation in L2 discourse: English tag questions used by Spanish speakers. *Intercultural Pragmatics*, 2 (2), pp.151-168.

Watson, C. & Cummins, J. 1995. *One language or two? Helping families from other cultures decide on how to talk to their language-delayed child*. Hanen Centre: WigWag.

Watt, D. 2002. 'I don't speak with a Geordie accent, I speak, like, the Northern accent': contact-induced levelling in the Tyneside vowel system. *Journal of Sociolinguistics*, 6 (1), pp.44-63.

Watterson, B. 1995. *The Calvin And Hobbes tenth anniversary book*. Kansas City, MO: Andrews and McMeel.

Wei, L. 2000. Dimensions of bilingualism. *The Bilingualism Reader*, 3, pp1-541.

Weismer, G. 1979. Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics*, 7 (2), pp.197-204.

Wells, J.G. 1962. *A study of the formants of the pure vowels of British English.* Doctoral dissertation. University of London.

Willems, N.J. 1982. *English intonation from a Dutch point of view*. Dordrecht: Foris Publications.

Winter, B. & Wieling, M. 2016. How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling. *Journal of Language Evolution*. [E-journal]. 1 (1), pp.7-18. Available at: http://dx.doi.org/10.1093/jole/lzv003.

Wu, C.H. & Shih, C. 2012. A corpus study of native and non-native vowel quality. *Proceedings of the 6th International Conference on Speech Prosody*. Shanghai, China, 22nd-25th May 2012. Shanghai: Tongji University Press, pp.234-237.

Xu, Y. 2013. Prosodypro a tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP). Aix-en-Provence, France. pp.7-10.

Yang, J. & Fox, R.A. 2017. L1–L2 interactions of vowel systems in young bilingual Mandarin-English children. *Journal of Phonetics*, 65, pp.60-76.

Yeni-Komshian, G.H., Garamazza, A. & Preston, M.S. 1977. A study of voicing in Lebanese Arabic. *Journal of Phonetics*, 5 (1), pp.35-48.

Zerbian, S. 2015. Markedness considerations in L2 prosodic focus and givenness marking. In: *Prosody and language in contact*. Springer, Berlin, Heidelberg. pp.7-27.

Zobl, H. 1980. The formal and developmental selectivity of L1 influence on L2 acquisition. *Language Learning*, 30, pp.43-57.

Appendices

Appendix A

a) The Arabic proficiency test (APT):

Example section one:

أولاً: الأصوات وفهم المسموع:

١- سماء

٢- سماح

٣- سماع

Sounds and understanding what you hear:

1- /sama:ʔ/ sky

2- /sama:ħ/allowing

3- /Sama:ʕ/ listen

Example section two:

اختر الإجابة الصحيحة:

١- لم أكن .... حصولي على المركز الأول في المسابقة.

أ/ أتوقع

ب/ أودّ

ج/ أعتقد

د/ أرى

Choose the correct answer:

1- I did not………to get the first place in the competition.

a/ expect

b/ want

c/ think

d/ see

Example section three:

اختر الإجابة الصحيحة:

١- قرأت ثلاثة ....

أ/ كتاب

ب/ كتب

ج/ الكتاب

Choose the correct answer:

1- I read three………

a/ book

b/ books

c/ the book

b) Examples from the online TOEFL test:

Example section one:

(Audio recording)
Man: My car is broken.
Woman: I'll pick you up.
Narrator: What does the woman imply?

(Written answer choices; correct answer in bold face)

- **She'll give the man a ride.**
- She'll take the take the car to a gas station.
- She'll look for a new car.
- She'll ask for help.

Example section two:

Listen to the conversation and answer the question.
(Audio recording about a conversation between two friends; the conversation ends with the question: What does Johan want to buy?)

(written answer choices; correct answer in bold face)

- A mug with the university logo
- **A T-shirt with a matching hat**
- A book
- A calendar

Example section three:

When did Mary ……….. college?

(Correct answer in bold face)
- graduate
- **graduate from**
- graduating
- graduating from

Appendix B

Background Questionnaire (Arabic-English bilinguals)

Part 1

Participant code:……………………….

Please answer the following questions:

1. Age:……………………………………..
2. Sex: ☐Male ☐Female
3. Education (degree obtained or school level attended):………………………………………
4. Please list all the places where you have lived:

   Town/City……………………Country………………from……………..until……......
   Town/City……………………Country………………from……………..until……......
   Town/City……………………Country………………from……………..until……......
   Town/City……………………Country………………from……………..until……......
   Town/City……………………Country………………from……………..until……......
   Town/City……………………Country………………from……………..until……......

5. Which languages do you speak? And how well do you speak them?

   Language 1:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

   Language 2:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

   Language 3:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

6. Which of the above languages is your native language? (If you grew up with more than one language, please specify)

……………………………………………………………………………………………..

Part 2

1. How did you learn English as a foreign language? Please check all the answers that apply.
   - ☐ Through lessons while you were still in Saudi/Yemen?
     From…………………………..until…………………………………………
   - ☐ Through systematic English lessons after you emigrated?
     From…………………………..until…………………………………………
   - ☐ Only through everyday use?
     From…………………………..until…………………………………………
   - ☐ Other……………………………………………………………………….
     ……………………………………………………………………………..
2. With whom did you talk Arabic after the time of your emigration? If you no longer speak Arabic with these people, please indicate until when you did so?
   - ☐ Other……………………………………………………………………….
     ……………………………………………………………………………..
3. With whom did you talk Arabic after the time of your moving here? If you no longer speak Arabic with these people, please indicate until when you did so?
   - With your parents

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your parents until……………….

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your brothers and sisters

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your brothers and sisters until…………

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your partner or spouse

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your partner or spouse until…………..

☐Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐Never 0%

- With your children

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children until……………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:……………………………………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:…………………… until…………..

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

4- With whom did you talk English after the time of your moving here? If you no longer speak English with these people, please indicate until when you did so?

- With your parents

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your parents until……………….

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters until…………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse until……………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children until……………

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:……………………………

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:………………………… until…………………..

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- What is the native language of your partner or spouse?

  ………………………………………………………………………………

  ………………………………………………………………………………

- If you had another partner or spouse at an earlier time: what was his/her native language?

  ………………………………………………………………………………

  ………………………………………………

- Did you talk Arabic with him/her (your previous partner or spouse)?

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

If so until what time?...

5- How is talking Arabic nowadays?

☐ effortless ☐ with some effort ☐ difficult

6- For what other purposes do you use Arabic now? Please check all the answers that apply.

☐ when writing your diary
☐ in creative writing
☐ when writing your memories
☐ when writing letters
☐ when dreaming
☐ never
☐ other:…………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………

7- Is there anything else you would like to tell me about how you feel about Arabic or English?

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
………………………………………………………………

Thank you ☺

Background Questionnaire (English-Arabic bilinguals)

Part 1

Participant code:………………………

Please answer the following questions:

1. Age:……………………………………..
2. Sex: ☐Male ☐Female
3. Education (degree obtained or school level attended):………………………………………
4. Please list all the places where you have lived:
   Town/City………………………Country………………from………………until……....
   Town/City………………………Country………………from………………until……....
   Town/City………………………Country………………from………………until……....
   Town/City………………………Country………………from………………until……....
   Town/City………………………Country………………from………………until……....
   Town/City………………………Country………………from………………until……....
5. Which languages do you speak? And how well do you speak them?
   Language 1:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

   Language 2:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

   Language 3:…………….
   Proficiency:
   ☐ beg. ☐ int. ☐ adv. ☐ flu.

6. Which of the above languages is your native language? (If you grew up with more than one language, please specify)
……………………………………………………………………………………………..

Part 2

1. How did you learn Arabic as a foreign language? Please check all the answers that apply.

- ☐ Through lessons while you were still in the UK?
  From……………………………until…………………………………………
- ☐ Through systematic Arabic lessons after you emigrated?
  From……………………………until…………………………………………
- ☐ Only through everyday use?
  From……………………………until…………………………………………
- ☐ Other…………………………………………………………………………
  ……………………………………………………………………………..

2. With whom did you talk English after the time of your emigration? If you no longer speak English with these people, please indicate until when you did so?

- ☐ Other…………………………………………………………………………
  ……………………………………………………………………………..

3. With whom did you talk English after the time of your moving here? If you no longer speak English with these people, please indicate until when you did so?

- With your parents

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your parents until………………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters until…………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse until……………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children until……………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:……………………………………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:…………………… until…………..

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

4- With whom did you talk Arabic after the time of your moving here? If you no longer speak Arabic with these people, please indicate until when you did so?

- With your parents

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your parents until……………….

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your brothers and sisters until…………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your partner or spouse until……………

☐ Always 100% ☐ Frequently 80% ☐Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With your children until……………

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:……………………………

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- With others, please say with whom:………………………… until…………………..

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

- What is the native language of your partner or spouse?
  ……………………………………………………………………………………
  ……………………………………………………………………………………
- If you had another partner or spouse at an earlier time: what was his/her native language?
  ……………………………………………………………………………………
  ………………………………………………
- Did you talk English with him/her (your previous partner or spouse)?

☐ Always 100% ☐ Frequently 80% ☐ Sometimes 50% ☐ Seldom 10% ☐ Never 0%

If so until what time?...

5- How is talking English nowadays?

☐ effortless ☐ with some effort ☐ difficult

6- For what other purposes do you use English c now? Please check all the answers that apply.

☐ when writing your diary
☐ in creative writing
☐ when writing your memories
☐ when writing letters
☐ when dreaming
☐ never
☐ other:…………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

7- Is there anything else you would like to tell me about how you feel about Arabic or English?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………

Thank you ☺

Appendix C

Examples of the mini-dialogues of *wh*-questions:

A- English

Please have a look at the following sentences and then read them 6 times:

Is this your ball?
No it is not.

Do you have a pen?
Yes, I have

Are they both helpful?
Yes they are.

What are you going to do about the roof?
We are going to repair it.

At what time did you leave Miami?
We left at 5:00p.m..

From where is your friend Linda leaving?
From London.

And what did you do in London?
We spent three wonderful weeks there.

And when are we landing?
I don't know.

On what day are we leaving?
Next Monday.

With which airline are we leaving?
Miami Airlines.

So what about your friend Marline?
She went home for Christmas.

From where are you coming this morning?
I am coming from Monaco.

And what is your maiden name?
My maiden name is Jones.

And when is the meeting?
Tomorrow.

From where is this wonderful lemon?
From M&S.

And who is your new friend Lora?
She is the girl who is new in town.

And where do we live?
We live in London.

Do you come from Greece?
No I don't

Is your last name Smith?
Yes it is.

What are you doing?
Nothing.

And Where is your house?
In the bay.

B- Arabic

انظر الى هذه الجمل ثم اقرأها ست مرات:

هل ستذهب الى المدرسه؟
نعم سأذهب.

ما هو اسمك الاول؟
اسمي ليلى.

هل تحتاج الى المساعدة؟
لا شكرا.

ومتى غادرت مكة؟
غادرت مكة الساعة الخامسة مساء.

اين هو المفتاح؟
على الطاولة.

و متى غادرت ميامي؟
غادرت في الساعة الخامسة مساء.

ومن اين ستغادر صديقتنا ليندا ليلا ؟
من لندن.

وماذا فعلت في لندن؟
قضينا ثلاث اسابيع جميلة.

و متى سوف نهبط؟
لا اعلم.

في أي يوم سوف نغادر؟
الاثنين القادم.

على أي خطوط جوية سوف نسافر؟
على خطوط ميامي الجوية.

وماذا عن صديقتك مريم؟
ذهبت لكي تقضي الاجازة مع عائلتها .

ومن اين انت يا ماجدة؟
من مكة المكرمة.

ومتى سيكون اللقاء مساء؟

الساعة العاشرة.

من اين اشتريت عصير الليمون؟
من السوق.

و من هي صديقتك الجديدة لورا؟
هي الفتاة الجديدة في البلدة.

و أين سوف نعيش؟
سنعيش في لندن.

هل ستسافر الى المانيا؟
نعم سأسافر.

اين هو القلم الاحمر؟
هو في الحقيبة.

و من هي لبنى؟
هي صديقتي.

و ما هو اسم عائلتك؟
اسم عائلتي الخالد.

Appendix D

Examples of the vowels production

A- English

Please read the following sentences <u>three times:</u>
1)
I said kid and then I left
I said bid and then I left
I said hid and then I left
2)
I said bad and then I left
I said cad and then I left
I said had and then I left
3)
I said put and then I left
I said could and then I left
I said hood and then I left

B- Arabic

الرجاء قراءة الجمل التاليه ثلاث مرات:(الرجاء الانتباه الى الحركات الفتحه، الضمه،الكسره)

1)

قُلت جَد و ثم ذَهبت .

قلت كَد و ثم ذهبت.

قلت هَد و ثم ذهبت.

2)

قلت ضِد وثم ذهبت.

قلت شِد وثم ذهبت.

قلت هِد وثم ذهبت.

3)

قلت قُد و ثم ذهبت.

قلت صُد و ثم ذهبت.

قلت هُد و ثم ذهبت.

4)

قلت قاد وثم ذهبت

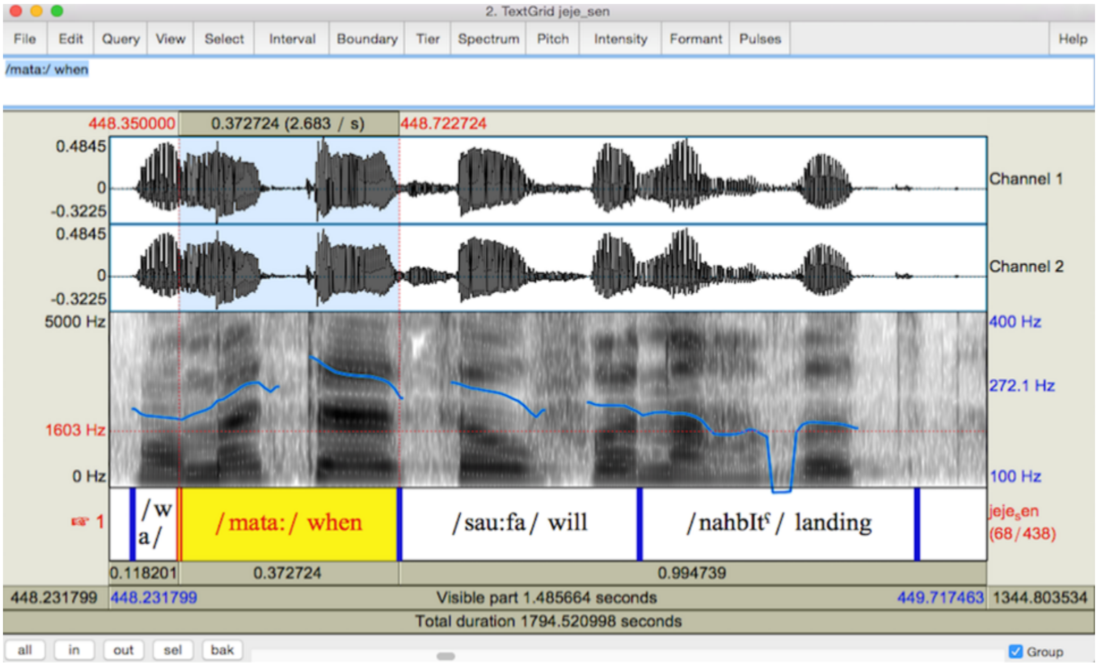
قلت صاد وثم ذهبت

قلت هاد وثم ذهبت

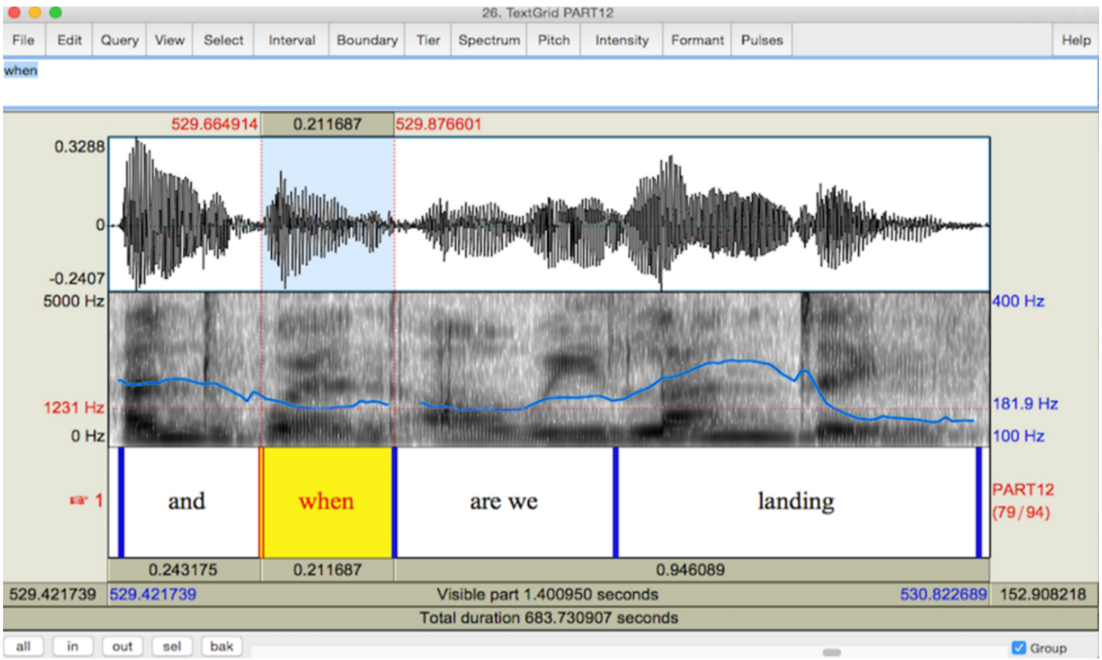
Appendix E

Example of Praat waveform and spectrogram:

A- Example: waveform and spectrogram of monolingual native Arabic speaker



B- Example: waveform and spectrogram of monolingual native English speaker



Appendix F

College of Arts & Humanities

Approval of Bangor ethics committee *of the College of Arts and Humanities*

PRIFYSGOL
BANGOR
UNIVERSITY

Ysgol Ieithyddiaeth ac Iaith Saesneg
Prifysgol Bangor

School of Linguistics and English Language
Bangor University

Myfyrwraig/*Student*: Alharbi, Amirah (LX-1430), July 8th 2015

Mae'r astudiaeth PhD hon wedi cael ei chadarnhau o ran agweddau moesegol, yn dilyn ymgynghoriad gyda'r arolygwr a swyddog Moeseg yr Ysgol, ac yn ogystal mae wedi cael ei chadarnhau gan bwyllgor Moeseg Coleg y Celfyddydau a'r Dyniaethau. Mae rhyddid i'r fyfyrwraig a enwir uchod barhau gyda chasglu'r data a gweithio ar y traethawd hir.
This PhD study has been approved with regards to ethical concerns, following consultation with the supervisor and the School Ethics officer, and furthermore has been approved by the Ethics committee of the College of Arts and Humanities. The student named above is now free to continue with collecting the data and working on the dissertation.

Dr Peredur Webb-Davies
Swyddog Moeseg yr Ysgol / Darlithydd mewn Ieithyddiaeth Gymraeg
School Ethics officer / Lecturer in Welsh Linguistic

PRI FYS GOL

APPLICATION FOR APPROVAL BY RESEARCH ETHICS COMMITTEE BANGOR
STUDENT RESEARCH PROJECT - DECISION UNIVERSITY

Reference Number: CAH 1 1

Name of researcher:	Amirah Alharbi
Name of supervisor:	Peredur-Webb Davies
Research project title:	First Language Attrition and Second Language Acquisition: The Effects of Phonetic Aptitude on Arabic-English Late Bilinguals
Title of qualification being pursued:	PhD Linguistics and English Language
Submission date:	20/05/1 5
Decision date:	1 9/08/1 5

This proposal was approved by College of Arts and Humanities Ethics Committee, Bangor University, on

Signed (on behalf of the committee) by:

Peter Shapely
Date: 1 9/08/1 5

Appendix G

A. Participant Consent Form English version:

PRIFYSGOL
BANGOR
UNIVERSITY

Bangor University's 'Code of Practice for the Assurance of Academic Quality and Standards of Research Programmes' (Code 03)
https://www.bangor.ac.uk/ar/main/regulations/home.htm

COLLEGE OF ARTS & HUMANITIES

Participant Consent Form

Researcher's name:

The researcher named above has briefed me to my satisfaction on the research for which I have volunteered. I understand that I have the right to withdraw from the research at any point. I also understand that my rights to anonymity and confidentiality will be respected.

I agree that my speech may be audio recorded and that the audio recordings will later be analysed as part of the research project. I also agree that the audio recordings may be used for future research projects.

Signature of participant ………………………………………………………………

Date ………………………………………………………………

This form will be produced in duplicate. One copy should be retained by the participant and the other by the researcher.

B. Participant Consent Form Amirah version

PRIFYSGOL
BANGOR
UNIVERSITY

"مدونة الممارسات لضمان الجودة الأكاديمية
ومعايير البرامج البحثية" لدى جامعة بانجور (المدونة 03)
https://www.bangor.ac.uk/ar/main/regulations/home.htm

كلية الآداب والعلوم الإنسانية

نموذج موافقة المشارك

اسم الباحث:

قدّم لي الباحث المذكور أعلاه موجزاً وافياً عن البحث الذي تطوّعت للمشاركة فيه. أفهم بأن لي الحق بالانسحاب من البحث في أية مرحلة، كما أفهم تماماً حقوقي باحترام السرية وكتمان هويتي.
أوافق على تسجيل حديثي صوتياً وعلى أن تلك السجلات الصوتية ستخضع لاحقاً للتحليل ضمن المشروع البحثي. كما أوافق على استخدام تلك التسجيلات الصوتية لأغراض المشاريع البحثية المستقبلية.

توقيع المشارك: ....

التاريخ: ....

سيتم إعداد نسختين من هذا النموذج بحيث يحتفظ المشارك بإحدى النسختين ويحتفظ الباحث بالنسخة الأخرى.

Appendix H

A. Participant Information Sheet English Version

PRIFYSGOL
BANGOR
UNIVERSITY

Bangor University's 'Code of Practice for the Assurance of Academic Quality and Standards of Research Programmes' (Code 03)
https://www.bangor.ac.uk/ar/main/regulations/home.htm

COLLEGE OF ARTS & HUMANITIES

PARTICIPANT INFORMATION SHEET:

L1 and L2 skills in highly proficient bilinguals

The School of Linguistics and English Language at Bangor University requires that all persons who participate in linguistic studies give their written consent to do so. Please read the following carefully before you decide whether or not to participate in the present study.

The broad goal of this research is to determine the relationship between the first language and the second language in highly proficient bilinguals. If you decide to participate, you will be asked to:

- fill in a background questionnaire (this will take around 10 minutes)
- do a grammar (structure) test and listening comprehension test (this will take around 30 minutes)
- do a spelling game, a word language task, and a word discrimination task (this will take around 60 minutes)
- describe carton pictures, produce words and sentences, and read sentences aloud (this will take around 30 minutes and your speech will be audio recorded)

You will do these tasks in two separate sessions of about 60 to 70 minutes.

Your responses and information from this study will be kept strictly confidential. Your name will not be associated in any way with the research findings. Data collected from this experiment may be used for future research or made available to other research staff working on the project, but all data will be preserved or distributed in a completely anonymous format

– with no personal identification marks.

Your participation is voluntary. You are free to withdraw consent from the study and discontinue participation at any time and for any reason without any negative consequences.

The data from this study will be used for the researcher's PhD dissertation. It may also be presented at conferences or as journal publications. In all cases, all data reported will be completely anonymous.

If you have any questions about the tasks or procedure of the study, please ask the researcher before signing the consent form. If all of your questions have been answered to your satisfaction and you would like to participate in the current study, we would ask you to sign the consent form.

If you have any general questions relating to the project, feel free to contact Amirah Saud A. Alharbi at elp2dd@bangor.ac.uk.

If you have any concerns or complaints about this study, please contact Dr. Marco Tamburelli, Head of School, School of Linguistics and English Language, Bangor University at the following address:

School of Linguistics and English Language
Bangor University
College Road
Bangor, Gwynedd
LL57 2DG

B. Participant Information Sheet Arabic Version

PRIFYSGOL
BANGOR
UNIVERSITY

"مدونة الممارسات لضمان الجودة الأكاديمية
ومعايير البرامج البحثية" لدى جامعة بانجور (المدونة 03)
https://www.bangor.ac.uk/ar/main/regulations/home.htm

كلية الآداب والعلوم الإنسانية

صفحة معلومات المشارك

مهارات اللغة الأولى واللغة الثانية لدى المتحدثين بطلاقة بلغتين

تطلب كلية اللغويات واللغة الإنجليزية في جامعة بانجور أن يقدم كافة المشاركين في دراسات اللغويات موافقتهم الخطية على المشاركة. نرجو منكم قراءة ما يلي بعناية قبل أن تقرروا ما إن كنتم ترغبون بالمشاركة في هذه الدراسة.

يتمثل الهدف الواسع من هذه الدراسة في تحديد العلاقة بين اللغة الأولى واللغة الثانية لدى المتحدثين بطلاقة بلغتين. في حال قررت المشاركة، سيطلب منك ما يلي:

- ملء استبيان بالمعلومات العامة (سيستغرق منك حوالي 10 دقائق)
- خوض اختبار (هيكلي) للقواعد واختبار للاستيعاب السمعي (سيستغرق ذلك حوالي 30 دقيقة)
- القيام بلعبة للتهجئة ومهمة للّغة والكلمات ومهمة أخرى لتمييز الكلمات (سيستغرق ذلك حوالي 60 دقيقة)
- وصف صور كرتونية وإنتاج الكلمات والجمل وقراءة الجمل بصوت مسموع (سيستغرق ذلك حوالي 30 دقيقة وسيتم تسجيل حديثك)

ستقوم بتلك المهام على جلستين منفصلتين يتراوح طول كل منهما بين 60 و 70 دقيقة.

ستتم مراعاة أقصى درجات الخصوصية لإجاباتك ومعلوماتك المتعلقة بهذه الدراسة، ولن يتم ربط اسمك بأي من نتائج البحث بأي طريقة. من الممكن أن يتم استخدام البيانات التي يجري الحصول عليها من خلال هذه التجربة في الأبحاث المستقبلية أو توفيرها لباحثين آخرين يعملون على المشروع، ولكن كافة البيانات ستحفظ أو توزع بشكل سري تماماً بحيث لا تحمل أي اسم أو معلومات شخصية تعرّف بصاحبها.

مشاركتك في هذه الدراسة طوعية، ويمكنك سحب موافقتك على المشاركة في الدراسة والتوقف عن المشاركة في أي وقت ولأي سبب دون أن يترتب على ذلك أية نتائج سلبية.

ستستخدم البيانات التي تخلص إليها هذه الدراسة في رسالة الدكتوراه التي يعدّها الباحث، كما قد يتم عرضها خلال المؤتمرات أو ضمن نشرات دورية. وفي كافة الأحوال، لن ترتبط البيانات بأية هوية أو معلومات تفصح عن صاحبها.

إذا كانت لديك أية أسئلة تتعلق بالمهام أو الإجراءات التي تنطوي عليها هذه الدراسة، نرجو منك توجيه أسئلتك للباحث قبل التوقيع على نموذج الموافقة. وفي حال الإجابة على كافة الأسئلة على نحو مُرضٍ، ورغبت في المشاركة في هذه الدراسة، نرجو منك التوقيع على نموذج الموافقة.

إن كانت لديك أسئلة عامة تتعلق بالمشروع، فنرحب بتواصلك مع أميرة سعود الحربي على العنوان الإلكتروني: elp2dd@bangor.ac.uk.

إذا كانت لديك أية مخاوف أو شكاوى تتعلق بهذه الدراسة، نرجو منك التواصل مع الدكتور ماركو تامبوريلي، عميد كلية اللغويات واللغة الإنجليزية في جامعة بانجور على العنوان التالي:

كلية اللغويات واللغة الإنجليزية

جامعة بانجور

كوليج رود،

بانجور

غوينيد

LL57 2DG

المملكة المتحدة

Appendix I

Approval consent from ALI, UQU

المملكة العربية السعودية
وزارة التعليم العالي
جامعة أم القرى

الرقم : ....
التاريخ : 18-5-2015
المشفوعات : ....

Dear Ameera Saud AlSaedi ALHarbi

(PhD student at Bangor University, UK, and Teaching Assistant at Umm al-Qura University),

In response to your letter requesting to implement your study on the students of the Arabic Language Institute, UQU, Makkah concerning your research entitled: <u>First Language Attrition and Second Language Acquisition in the Domain of Phonetics: The Impact of Phonetic Aptitude and Bilingualism</u>, let me assure you of my consent to conduct this study and implement it on our students, males and females, wishing you all the best.

Dr. Adel Ahmad Banaemah

Dean, 18/5/2015

Arabic Language Institute

Umm Al Qura University
Makkah Al Mukarramah P.O. Box: 715
Cable Gameat Umm Al- Qura, Makkah
Faxemely: 02 - 5564560 \ 02 - 5593997
Tel Aziziyah: 02-5501000 Abdiyah: 02 - 5270000

جامعة أم القرى
مكة المكرمة ص.ب: ٧١٥
برقيا: جامعة أم القرى – مكة
فاكسميلي: ٠٢-٥٥٦٤٥٦٠ / ٠٢-٥٥٩٣٩٩٧
تلفون سنترال العزيزية: ٠٢-٥٥٠١٠٠٠ العابدية: ٠٢-٥٢٧٠٠٠٠