University of Central Lancashire
UCLan

The L2 Perception and Production of English Vowels by L1 Greek-Cypriot Speakers

Dimitra Dimitriou

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

to

School of Humanities, Language & Global Studies

University of Central Lancashire

September 2023

RESEARCH STUDENT DECLARATION FORM

Type of Award **Doctor of Philosophy**_____

School **Humanities, Language & Global Studies**_____

*Sections marked * delete as appropriate*

1. Concurrent registration for two or more academic awards

*I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

2. Material submitted for another award

***I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work

3. Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

N/A_____

4. Use of a Proof-reader

*No proof-reading service was used in the compilation of this thesis.

Signature of Candidate _____

Print name: Dimitra Dimitriou_____

Abstract

Previous research shows that acquiring L2 segments is one of the most challenging tasks L2 learners face, particularly when the L1 and L2 inventories involve different contrasts and acoustic cues. The present research investigated the perception and production of L2 English vowels by adult Greek-Cypriot learners and examined the effects of High Variability Phonetic Training (HVPT) on their perceptual and production performance, as well as learners' ability to generalise new knowledge to new speakers and contexts and retain any improvement two months later. Individual differences in motivation, input and language use patterns were also assessed. The HVPT protocol followed included 8 training sessions with 330 natural and synthetic stimuli each, involving both real and non-words. Perceptual performance was assessed through a forced-choice identification task and production performance through a wordlist-reading and an elicited imitation task, both analysed acoustically and through intelligibility ratings by native English speakers. Individual factors were assessed using a questionnaire complemented by qualitative, semi-structured interviews. The findings clearly showed the influence of the L1 on both the perception and production of L2 segments, supporting the assumptions of current models of speech perception and production. Learners faced challenges in perceiving the members of L2 contrasts and mostly used their L1 articulatory routines in their productions of L2 vowels. The HVPT protocol followed was found to be insufficient to lead to significant improvements in overall perceptual or production performance, although some improvements were observed in either perceiving or producing some target vowels. Individual learner differences had a significant effect on participants' performance suggesting that they merit more attention than they currently receive. This study was the first to provide an in-depth examination of the acquisition of L2 English vowels by Greek-Cypriot learners and the factors that may affect their performance, thereby guiding future research as well as EFL practitioners.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to thank my supervisory team, Dr Antri Kanikli, Dr Daniel Waller and Dr Daniel Matthias Bürkle, for their tireless support, advice and guidance. They have been there for every minor or major problem, every moment of doubt but also success, encouraging me and cheering for me in every step of the way. I could not have asked for a better team, and I consider myself lucky to have been able to work with them – I can only hope that they feel the same way! I would also like to thank Dr Elena Kkese, who was part of this team for a brief period of time, but has happily provided her knowledge and support long after as well. Her comments and advice were invaluable for the completion of this thesis. Dr Hae-Sung Jeon's comments on earlier drafts have also been insightful, and I am truly thankful for the opportunities she offered me for training and learning.

Of course, I am immensely grateful for the support of my family as well: my mum, who listened to me describe, explain, justify and complain for endless hours, offering only words of encouragement and pride; my dad, the silent supporter in difficulties and achievements who was always encouraging me to remember to "take a break" and "go on vacation"; my (now) fiancé, who has been there since the very beginning and has been patient, supportive and understanding along this very (very) long journey! I should also thank my dear friends and siblings, who are still waiting for me to make time for a coffee, for their support and patience! I knew I could always count on them, and I promise I will make up for the lost time!

My sincerest thanks also go to all the people who have participated in this study, sometimes spending hours of their time to complete the tasks! I would like to thank the 10 native English speakers who provided their recordings, the 18 Greek-Cypriots who started and/or completed the tasks, and in particular the 8 Greek-Cypriots who also spent the time to complete the training sessions, the 2 phoneticians and 3 naïve listeners who rated the recordings of native English speakers and the 5 naïve listeners who assessed the recordings of Greek-Cypriot speakers as well as the 10 volunteers who helped in pilot-testing the questionnaires.

Last but not least, I am sincerely thankful for the financial support provided to me by the A. G. Leventis Foundation and by the Arts and Humanities Research Academy of the University of Central Lancashire.

To Danae

Abbreviations

AOA	Age of Arrival
AOL	Age of Learning
CAPT	Computer-Assisted Pronunciation Training
CG	Category-Goodness
CYG	Cypriot-Greek
DV	Dependent Variable
EFL	English as a Foreign Language
ESL	English as a Second Language
FCID	Forced-Choice Identification Task
FL	Functional Load
FLA	Foreign Language Acquisition
GA	General American
HVPT	High Variability Phonetic Training
IV	Independent Variable
L1	Native Language
L2	Second Language
LOR	Length of Residence
LVPT	Low Variability Phonetic Training
NA	Non-Assimilated
NE	Native English
PAM	Perceptual Assimilation Model
RP	Received Pronunciation
RQ	Research Question
SC	Single-Category
SLA	Second Language Acquisition
SLM	Speech Learning Model
SMG	Standard Modern Greek
SSBE	Standard Southern British English
TC	Two-Category
TP	Teste/Treino de Percepção
UC	Uncategorised-Categorised
UU	Uncategorised-Uncategorised

CHAPTER 1: INTRODUCTION

Research in the field of Second and Foreign Language Acquisition (SLA and FLA, respectively) demonstrates that learners tend to retain a foreign accent in their second language (L2), irrespective of their desire to sound native-like (Suzukida & Saito, 2019). Learning to accurately perceive and produce new phonetic categories has been identified as one of the most challenging tasks L2 learners face. Adult L2 learners have been found to face difficulties in the perception and production of both L2 consonants (e.g. Best, McRoberts, & Goodell, 2001; Dimitriou, 2019; Hattori & Iverson, 2009, 2010; Iverson et al., 2003) and vowels (e.g. Flege & MacKay, 2004; Iverson & Evans, 2007; Lengeris, 2009a). More specifically, L2 learners have been found to encounter difficulties in distinguishing the sounds of an L2 contrast when they are not contrastive in their native language (L1), due to the fact that they are less accustomed to attending to the acoustic cues that native speakers use for the reliable discrimination of the sounds in the contrast (Aliaga-Garcia & Mora, 2009; Cebrian, 2006, 2007; Escudero, 2005; Flege, Frieda, & Nozawa, 1997a; Kkese & Petinou, 2017; Kondaurova & Francis, 2010; Mora & Fullana, 2007; Morrison, 2003; Zhi & Li, 2021).

The difficulties that adults face when learning L2 vowels are even greater when the L1 and L2 vowel inventories are different, particularly when the L1 vowel inventory is smaller than the L2 vowel inventory (e.g. Balas, 2018; Bogacka, 2004; Georgiou, 2019; Jin & Liu, 2014; Lengeris & Hazan, 2010; Lengeris, 2009a; Rojczyk, 2010; Yuan & Archibald, 2022; Zhi & Li, 2021). More specifically, when a single L1 vowel category exists in the acoustic space occupied by two or more L2 vowels, it is likely that the two unfamiliar sounds of the L2 will be perceived as exemplars of the same category by the learners. For example, it was observed that Spanish learners of English, whose L1 only has five vowels compared to the greater number of vowels in the inventory of their L2, have difficulties in discriminating English vowel contrasts (Cebrian, 2006; Flege, Bohn, & Jang, 1997b; Iverson & Evans, 2009; Sakai, 2016).

Similar problems in distinguishing English vowel contrasts were observed for Greek learners of English as well (Lengeris, 2009a). The phonemic inventories of Standard Modern Greek (SMG) and Cypriot-Greek (CYG) are considerably different from the phonemic inventory of English, especially in terms of their vowel systems, which differ both in size and complexity (Coutsougera, 2007). Given that CYG has a smaller vowel inventory than English and does not use spectral or durational cues to signal phonological contrast, it is expected that these

learners of English will have difficulties in successfully using these cues to detect the subtle differences between contrasting vowels which overlap a single L1 category (Best & Tyler, 2007; Best, 1995; Flege, 1995). As a result of this, learners are likely to produce contrasting vowels without the required durational or spectral differences, merging them in a single, L1-based category.

Importantly, great individual differences have also been observed in the perceptual and production abilities of learners in novel L2 speech sounds, even among learners who share the same L1 background and similar profiles (Lengeris & Hazan, 2010; Lengeris, 2009a; Munro, Derwing, & Thomson, 2015; Piske, MacKay, & Flege, 2001; Qian, Chukharev-Hudilainen, & Levis, 2018). Various factors have been argued to influence the success of L2 phonetic learning in addition to the relation between the segmental inventories of the L1 and the L2, among which the age of learning (AOL) of the L2 (e.g. Flege, Yeni-Komshian, & Liu, 1999a), the length of residence (LOR) in an L2 setting (e.g. Flege & Liu, 2001) and L1-L2 use patterns (e.g. Flege & MacKay, 2004; Flege et al., 1997a; Piske et al., 2001). Other factors that may have a role to play in L2 phonetic learning include motivation to learn (e.g. Bongaerts, van Summeren, Planken, & Schils, 1997; Moyer, 1999), and quality and quantity of input (e.g. Flege & Liu, 2001; Flege, 2008).

Based on the assumption that the quality of the L2 experience is an important factor in overcoming age-related effects on degree of foreign accent (Flege et al., 1997b; Flege, Munro, & MacKay, 1995; Thomson, 2018), it has been argued that phonetic training can compensate for the lack of the naturalistic exposure required for the formation of L2 phonetic categories in foreign language classrooms, by providing targeted input that can help learners better perceive problematic phonemes (Sakai & Moorman, 2018). Previous research has demonstrated that Computer-Assisted Pronunciation Training (CAPT), especially when it includes highly variable stimuli with the target sounds in multiple phonetic environments, is effective in promoting segmental perception (e.g. Nishi & Kewley-Port, 2007, 2008; Thomson, 2011; Wang & Munro, 2004, among others).

This has led to the development of High Variability Phonetic Training (HVPT), a technique that can be incorporated in CAPT to make pronunciation instruction more effective in achieving measurable perceptual changes for adult L2 learners, as it provides exposure to highly variable stimuli (Thomson, 2011, 2018; Wang & Munro, 2004). According to Thomson

(2018, p. 220), “in its original and most basic form, HVPT uses natural training stimuli produced by multiple talkers, in multiple phonetic contexts”, thereby providing exposure to numerous samples of the auditory training stimuli, as opposed to what is the current practice in the language classroom.

Based on the findings of previous studies as outlined above, the current research aims to investigate the following Research Questions (RQs):

1. How do CYG adult learners perceive and produce L2 English vowels?
2. What are the spectral and durational differences in English vowel production between CYG learners and native English (NE) speakers?
3. How effective is HVPT in improving CYG learners’ vowel perception and production without explicit production training?
4. How well does any improvement generalise to new speakers and contexts, and to what extent is it retained after a two-month period?
5. Do individual differences in motivation, input and language use patterns affect CYG learners’ perception and production of L2 English vowels?

This study focuses on the acquisition of vowels not only because the two languages differ substantially in that respect, but also because it is argued that "L2 vowels are more difficult to learn (…) regardless of the learners' L1 background" (Thomson, 2011, p. 750). Furthermore, vowels were found to contribute more to the intelligibility of utterances than consonants (Bent, Bradlow, & Smith, 2007), while it was also observed that speakers’ intelligibility of L2 consonants may improve over time without intervention, as opposed to L2 vowels (Neri, Cucchiarini, & Strik, 2006).

The thesis is structured as follows: Chapter 2 presents a review of relevant literature, followed by a detailed description of the methodology followed (Chapter 3) and the results obtained through quantitative and qualitative analysis (Chapter 4). Chapter 5 discusses these findings while Chapter 6 presents some concluding remarks.

The following chapter offers more details about the current linguistic situation in Cyprus (2.1) as well as the characteristics of learning in an EFL context (2.2). Section 2.3 focuses on the acoustic characteristics of the vowels in SMG and the CYG dialect compared to the vowels

in English, followed by a discussion of the orthographic systems of the two languages (2.4). Section 2.5 deals with current models of L2 perception and production, the relationship between the two modalities and the influence of the L1 on L2 production and perception focusing on the perceptual assimilation patterns of English vowels onto CYG vowels. Section 2.6 presents some factors that may affect L2 segmental acquisition and degree of foreign accent, i.e. language use patterns, motivation and input. The notions of intelligibility and comprehensibility and the implications of foreign-accentedness in L2 speech are discussed in 2.7. Finally, 2.8 presents the HVPT technique along with its advantages and applications in previous research, and 2.9 discusses some methodological considerations when implementing the HVPT paradigm.

CHAPTER 2: LITERATURE REVIEW

2.1 The Linguistic Situation in Cyprus

CYG is a dialect of SMG spoken mainly in Cyprus, but also by the immigrant communities in the UK, North America and Australia (Arvaniti, 1999a). CYG is spoken by approximately 1 million speakers, compared to approximately 15 million speakers of SMG (Simaki, Mporas, & Kondyli, 2015).

According to Simaki et al. (2015), the development of CYG as it is today has been influenced by both English and SMG, as can be explained by the historical background of the island. Cyprus was historically populated by Greeks since the Bronze Age, until Turks began settling on the island when it became part of the Ottoman Empire in 1571 (Arvaniti, 2006). In 1878, the control of Cyprus was handed to Britain, and the island became a British colony in 1925 until 1960, when Cyprus gained its independence (Yazgin, 2007). However, a series of clashes between the two ethnic groups that populated the island at the time, i.e. the Greek majority and the Turkish minority, led to a military invasion of the island by Turkey and the occupation of the northern third of Cyprus up to the present (Arvaniti, 2006). The majority of the Turkish-Cypriots now live in the northern part of the island, whereas the majority of the Greek-Cypriots live in the southern, non-occupied areas of the Republic of Cyprus (Arvaniti, 2006). The term CYG in this study refers to the variety of Greek spoken by the Greek-Cypriots in the southern part of the island, where it is the dominant language (Georgiou, 2019), whereas the term SMG refers to the standard variety spoken in Greece[1]. Although Turkish is spoken by only a small percentage of Greek-Cypriots in the non-occupied areas at present, the official languages in Cyprus are both Turkish and Greek, and they are both retained in official documents (Arvaniti, 2006; Georgiou, 2019).

The importance of the English language on the island, however, also merits attention. Even though the status of English changed from Second to Foreign language after the island became independent, the language retains its prominence in Cyprus, and it is commonly used in international communication and as a lingua franca both in social and professional settings

[1] For a comprehensive description of the linguistic features of CYG and a comparison with SMG, as well as a more detailed background of the current linguistic situation in Cyprus, its educational system and its history, the reader is directed to Arvaniti (2006).

(McEntee-Atalianis, 2004). For example, even though it is not an official language, English is still widely used in a large proportion of official documents in many domains, including administration, banking and healthcare, while it was also the exclusive language of the law until 1987 (Arvaniti, 2006; McEntee-Atalianis, 2004; Yazgin, 2007). English as a Foreign Language is also a compulsory part of primary and secondary education. More specifically, English was a compulsory lesson for students aged 9 to 16, and optional until 18. This was recently changed to be included as a compulsory lesson from the age of 6. Furthermore, the vast majority of CYG students attend afternoon lessons to obtain a qualification, most often an IGCSE (either by Edexcel or Cambridge International Examinations) or an IELTS certificate. English is also the medium of instruction in the majority of private secondary schools in Cyprus and of private tertiary colleges (McEntee-Atalianis, 2004).

Furthermore, English is regularly used in the everyday life of Greek-Cypriots, especially due to the island's status as a tourist destination requiring locals to use English to communicate with tourists (Yazgin, 2007). In addition, as a result of migration from Cyprus to the UK after the time of independence, many Greek-Cypriots have relatives in the UK, especially London, who are either living there permanently, or studying. This increases Greek-Cypriots' motivation to learn English, especially since it is viewed as a means to access universities in the UK or globally (Yazgin, 2007). English is also considered to be an essential skill for some forms of employment or some higher education institutions, while it is also viewed as a means to access information and travel (McEntee-Atalianis & Pouloukas, 2001; McEntee-Atalianis, 2004; Yazgin, 2007). All of these factors have contributed to the development of a positive attitude towards learning and using English in Cyprus, despite the fact that this was the colonisers' language (Yazgin, 2007).

In terms of the native varieties, although both CYG and SMG are used in Cyprus, there is a functional differentiation between them. More specifically, CYG is the variety acquired at home and used in everyday interactions among Cypriots, especially in oral conversation (Arvaniti, 1999a, 2006; Georgiou, 2019; McEntee-Atalianis, 2004; Simaki et al., 2015). SMG on the other hand is learnt through formal education, and it is used in most forms of writing and some forms of oral discourse (e.g. news broadcasting), although it is preferred in formal and semi-formal contexts (Arvaniti, 1999b, 2006; Georgiou, 2019; McEntee-Atalianis, 2004; Simaki et al., 2015). Arvaniti (2006), thus, argues that the linguistic situation in Cyprus can be

described as diglossic as per Ferguson's (1959) description, as CYG speakers consider the two varieties as distinct and there is agreement as to which circumstances require the use of each.

As a result of the general assumption that the variety of Greek used in Cyprus and the standard variety spoken in Greece are similar, the CYG dialect is generally neglected in the literature. Arvaniti (2006), however, presents data demonstrating that the variety used in Cyprus has been increasingly diverging from the standard, making the two varieties considerably different from each other. More specifically, even though CYG is described as a dialect of SMG, the two varieties are different to such an extent that SMG speakers often find CYG unintelligible, unless it contains some characteristics of accommodation to the SMG listener (Arvaniti, 1999a, 2006; Simaki et al., 2015). CYG speakers, on the other hand, can understand SMG more easily, given that it is the official language of Cyprus, used in education and the media (Arvaniti, 1999a, 2006).

Furthermore, the fact that CYG is mostly used in conversational speech reinforces the belief that CYG is just a regional accent, as its pronunciation is considered to be one of the most distinctive characteristics differentiating it from SMG (Arvaniti, 2006). However, the differences between CYG and SMG are found in all levels of linguistic structure, both in oral and written discourse (Arvaniti, 2006; Simaki et al., 2015), and mostly in the lexicon, since each variety contains loan words from different sources. More specifically, CYG may use loans or literal translations from English for items that SMG has borrowed from French (e.g. [fail] "file" instead of the SMG [dosʲe] and [ham] "ham" instead of SMG [zambon]) (Arvaniti, 2006). CYG terms for everyday items may also be replaced by English loans: e.g. [antenna] "antenna" instead of SMG [kerea], [klip] "paper clip" instead of SMG [sinðetiras], reflecting the influence of English on the island (Arvaniti, 2006; Yazgin, 2007).

The following section will focus on the English language classroom in Cyprus, outlining the characteristics of learning in an EFL context.

2.2 The EFL Context

L2 learners in an immersion context enjoy naturalistic exposure that indisputably offers many benefits that can play a key role in the development of L2 speech. Among these are the opportunity to interact with L2 speakers with variable speaker characteristics (e.g. linguistic variability due to age, gender, dialect), thereby increasing the possibility of developing robust

L2 production and perception routines, and the exposure to a combination of both auditory and visual information, which has been found to be more beneficial in L2 speech production and perception than auditory information alone (e.g. Hazan et al., 2006; Hazan, Sennema, Iba, & Faulkner, 2005; Inceoglu, 2016).

This is not applicable to the FLA context, in which the L2 is mostly acquired through formal instruction largely confined to the L2 classroom (Georgiou, 2019). As noted by Best and Tyler (2007), FLA has the following characteristics: 1) the L2 is not widely used, 2) it does not extend significantly outside the classroom, 3) the emphasis is generally on formal instruction focused on vocabulary and grammar rather than on live conversation, and 4) the source of the L2 input is either L1-accented speech, or at best, speech by native L2 speakers using diverse L2 varieties, and thus learners are confronted with an incorrect or variable model of L2 phonetic details. Therefore, formal classroom instruction does not provide ideal conditions for L2 speech learning.

Since the input learners receive in a foreign language classroom is typically provided by non-native speakers, sometimes with noticeable foreign accents, or by a limited number of voices with clear pronunciation, providing little opportunity for learners to encounter native-like input (Barriusso & Hayes-Harb, 2018; Fabra & Romero, 2012; Georgiou, 2019; Hutchinson & Dmitrieva, 2022), it is often insufficient to prepare learners for the speaker variation or the variety of listening conditions that they may encounter (Crosby, 2020). It should be noted that even though native-speaker input is crucial when it comes to target-like perception and production of L2 speech, Tyler (2019) argues that foreign-accented speech in L2 interactions is not necessarily problematic, provided that phonological distinctions between all L2 phonemes are maintained and are perceived by native speakers as intended. However, when the input fails to provide clear phonological distinctions between L2 categories, as is often the case, learners are less likely to acquire them (Fabra & Romero, 2012; Hutchinson & Dmitrieva, 2022; Tyler, 2019).

Despite the increasing evidence that L2 pronunciation instruction can be beneficial for learners' L2 speech performance (see Derwing, Munro, Foote, Waugh, & Fleming, 2014; Krzonowski, Ferragne, & Pellegrino, 2015, among others), as well as the fact that pronunciation errors may have negative implications for L2 learners (Baran-Łucarz, 2011; Munro, 2003) (see 2.7), L2 pronunciation instruction is a highly neglected area in most ESL

and EFL classrooms, where training concerning the perception and production of L2 sounds is little to none, thereby preventing the acquisition of a more native-like pronunciation (Derwing & Munro, 2005; Georgiou, 2019; Lengeris, 2018; Papachristou, 2011; Piske et al., 2001). Pronunciation teaching is rarely included in the curriculum, and even when it is, many teachers either believe that improvement is impossible, or do not have the confidence or ability to teach it (Breitkreutz, Derwing, & Rossiter, 2001; Gilakjani, 2012; Lengeris, 2018). Therefore, students are very rarely provided with information about differences between phonemes in the L2, which they have to attain on their own (Gilakjani, 2012).

This lack of attention to pronunciation instruction applies to Cyprus as well, especially in state schools, where language teachers neglect pronunciation teaching mainly because it is considered to be a difficult task, there is a lack of teacher training and official guidelines in the curriculum, and teachers do not consider the cultivation of this skill important (Kyprianou, 2007, 2015). More specifically, Kyprianou (2007) found that despite the fact that textbooks include pronunciation activities, most teachers in Cyprus tend to neglect this skill and consider it to be the least important language skill to teach compared to others (83.6% of teachers reported skipping pronunciation activities; Kyprianou, 2007, p. 3). The few tasks practised in the classroom include listening to audio and video materials of native speakers and chorused imitation of the teacher, while no individual laboratory work or explicit phonetic training is provided. No instruction of phonetic symbols takes place in the classroom either, despite their existence in the textbooks (Kyprianou, 2007). Furthermore, pronunciation instruction tasks are not included in the curriculum, and they are not obligatory by the Ministry of Education (Kyprianou, 2007). In addition, teachers are not provided with specific guidelines or training by the Ministry on teaching pronunciation, even though they are trained twice a year on issues such as teaching writing or speaking.

This lack of attention to pronunciation and other skills related to communicative competence, in combination with the exposure to mostly foreign-accented speech, is likely to inhibit learning and prevent CYG learners from attuning their L1 phonetic categories to the acoustic properties of the L2 sounds. Furthermore, the strongest influence on CYG learners' pronunciation is arguably exerted by the American-accented input encountered through movies, music and the media, similarly to the Serbian students in Čubrović and Bjelaković (2020), who reported that they preferred using the General American (GA) model in their

speech, possibly because they are exposed to this variety much more than to standard Southern British English.

The following sections provide more details about the acoustic characteristics of the vowels in SMG, CYG and English, highlighting the similarities and differences in the vowel systems of each language.

2.3 Vowel Systems

2.3.1 Standard Modern Greek (SMG) and Cypriot-Greek (CYG)

The SMG vowel system consists of five vowels /i, e, a, o, u/ and does not have tense-lax or short-long distinctions (Arvaniti, 1999b, 2007; Baltazani, 2007; Fourakis, Botinis, & Katsaiti, 1999; Koutsoudas & Koutsoudas, 1962; Lengeris, 2009a; Mackridge, 1985; Nicolaidis, 2003; Sfakianaki, 2002). The vowel system of CYG is very similar, containing the same five vowel qualities, although "there are some small but noticeable differences in terms of the position they occupy in the vowel space" (Georgiou, 2019, p. 4; see also Arvaniti, 1999a; Themistocleous & Logotheti, 2016; Themistocleous, 2017a, 2017b).

Vowel spectral characteristics are typically measured through formant values: the first (F1) is inversely related to vowel height (higher F1 indicates lower vowel), the second (F2) is related to the vowel's backness (higher F2 indicates fronter vowel) and the third (F3) to the vowel's roundness (higher F3 indicates less lip-rounding). According to Fourakis et al. (1999), the size of a vowel inventory affects the position of the vowels in the F1xF2 acoustic space, as shown by Jongman, Fourakis, and Sereno (1989), who found that the SMG stressed vowels are "well separated in an acoustic space, allowing for maximal contrast between vowel categories" (Fourakis et al., 1999, p. 29).

However, it should be noted that the actual acoustic characteristics of vowels may be influenced by various factors, including segmental environment and speaker, as well as suprasegmental factors such as lexical stress, focus and speaking rate (Fourakis et al., 1999; Nicolaidis, 2003). For example, in an examination of the acoustic variability of SMG in spontaneous speech, Nicolaidis (2003) found evidence of overlapping formant distributions, which suggested that there was a lack of clear differentiation between the vowel categories in the F1xF2 acoustic space. More specifically, the author reported extensive overlap in the centre of the vowel space resulting in less differentiated vowels in spontaneous speech. The author also reported an upward shift of the vowel space in the fast rate condition, and a

movement of low and mid vowels towards a more centralised position in the vowel space, resulting in extensive overlap between /e/, /o/ and /a/ at the centre of the vowel space (Nicolaidis, 2003). Similar results were obtained by Fourakis et al. (1999), who examined the acoustic characteristics of SMG vowels in slow and fast tempo and in stressed and unstressed syllables. The researchers found that unstressed vowels have lower F1 values and are more central than stressed vowels, meaning that in the unstressed condition, the vowel system shrinks and shifts in a higher position in the F1xF2 acoustic space. According to Fourakis et al. (1999), a decrease in duration as a result of faster speaking rate or a lack of stress changes the formant patterns of vowels, making them more centralised than the original vowel.

As concerns vowel duration in SMG, Fourakis et al. (1999) found the low vowel (/a/) to be the longest and the high vowels (/i/ and /u/) to be the shortest, in accordance with previous studies for SMG vowels (e.g. Dauer, 1980; Fourakis, 1986). Nicolaidis (2003) also found the same pattern (/a > e > o > i > u/), noting that vowels differed significantly in their mean durations. However, these studies disagree as to the durations of the two high vowels. While Nicolaidis (2003) found /i/ to be longer than /u/, Dauer (1980) and Fourakis (1986) found that the two vowels are very similar in duration, and Fourakis et al. (1999) found /u/ to be longer than /i/. This may be due to the difference in the number of stimuli analysed in these studies, or their use of different means for acoustic measurements.

Nicolaidis (2003) also found that the duration of SMG vowels increases significantly as an effect of stress, as reported by Themistocleous and Logotheti (2016) as well, who found that in the stressed condition, back vowels were longer than front vowels, and low vowels were longer than high vowels (/a > o > u > e > i/), while in the unstressed condition, low vowels were longer than high vowels (/a > o > e > u > i/). In addition, segments may be reduced in spontaneous speech, varying from very reduced (hypo-forms) to non-reduced (hyper-forms), in the efforts of speakers to balance successful communication and articulatory economy (Nicolaidis, 2003).

Similarly to English (see 2.3.2), context effects in SMG vowels were also observed in Nicolaidis (2003), mainly from the preceding consonant, indicating that there is carryover rather than anticipatory influence on the vowel midpoint. In addition, context effects were mainly observed in the F2 compared to the F1 frequency. This suggests that contextual effects occur along the front-back axis (Nicolaidis, 2003). As concerns the F1, context effects were

more limited, but were again more evident from the preceding rather than the following consonant, with the exception of /a/ (Nicolaidis, 2003).

Themistocleous and Logotheti (2016) and Themistocleous (2017a, 2017b) examined and compared the acoustic characteristics of the vowels in the two varieties, providing a first examination of the acoustic structure of CYG vowels compared to SMG vowels. Similarly to Fourakis et al. (1999), Themistocleous and Logotheti (2016) and Themistocleous (2017a) also found that vowels are more centralised when unstressed and more peripheral when stressed, in both varieties. Overall, the F1xF2 vowel spaces of SMG and CYG were found to share a lot of similarities, especially in stressed vowels, although there were some differences particularly in unstressed vowels (Themistocleous & Logotheti, 2016; Themistocleous, 2017a). One difference between the two varieties was the F1xF2 position of unstressed vowels, which were found to be lower in CYG than in SMG (Themistocleous & Logotheti, 2016). Similar results were obtained by Themistocleous (2017a), who found that SMG unstressed /i, a, u/ vowels were more raised than the corresponding CYG vowels. Furthermore, Themistocleous (2017a, 2017b) report that the two varieties are different in F3 values, with significantly lower F3 values in SMG /i o u/ compared to the corresponding CYG vowels. This means that SMG vowels involve more lip-rounding than the corresponding CYG vowels.

In addition, vowel duration was found to differ in the two varieties (Themistocleous & Logotheti, 2016; Themistocleous, 2017a, 2017b). Vowels in both varieties were significantly affected by stress, with stressed vowels being longer than unstressed vowels. However, as opposed to the findings reported for SMG vowel duration, low vowels in CYG were longer than high vowels (/a > o > e > u > i/) in both stress conditions. Overall, CYG vowels are longer than SMG vowels, but the largest differences in vowel duration between the two varieties were found in unstressed vowels, with CYG unstressed vowels being significantly longer than unstressed SMG vowels (Themistocleous & Logotheti, 2016; Themistocleous, 2017a).

The durational differences between the two varieties may provide a possible explanation for this, as suggested by Arvaniti (2001), who examined the effect of speaking rate on segmental duration in CYG and SMG. More specifically, Arvaniti (2001) found that in SMG, faster speaking rates led to a reduction in the vowel duration, whereas in CYG both vowels and consonants were reduced to similar extents, indicating that CYG vowels may exhibit less

reduction than SMG vowels (see also Loukina, 2009 for a comparison of vowel and consonant reduction in SMG, CYG and Thessalian Greek).

2.3.2 English

The vowel system of English is larger than the SMG and CYG vowel systems, and it exhibits regional variation (Hughes, Trudgill, & Watt, 2013; Watson, 2009; Wright, 1996). More specifically, English has a complex vowel system consisting of over twenty vowels, including monophthongs with length distinctions and diphthongs (Ladefoged & Johnson, 2011). In terms of monophthongs, Standard Southern British English (SSBE), i.e. the standard variety typically spoken in the southern part of England, which according to Deterding (1997) is similar to Received Pronunciation (RP), contains five tense (/iː, uː, ɜː, ɔː, ɑː/) and six lax (/ɪ, ʊ, e, æ, ʌ, ɒ/) vowels (Bohn & Steinlen, 2003; Deterding, 1997; Katamba, 2009; Roach, 2004), as well as the allophone /ə/ (Kkese & Petinou, 2017). Although tense vowels are typically phonetically longer than their lax counterparts (Leung, Jongman, Wang, & Sereno, 2016), it is their spectral properties that make these vowels phonemically contrastive (Yuan & Archibald, 2022). Therefore, the primary acoustic cue used to differentiate English tense and lax vowels is spectral quality, while duration is a secondary cue used in some varieties (Hillenbrand, Clark, & Houde, 2000; Rato & Carlet, 2020).

As a result, SSBE makes use of more space in the F1xF2 vowel space, as its vowels are more dispersed, and it has short-long vowel distinctions that are lacking in the Greek varieties. All vowels can occur in stressed syllables, except for /ə/, which can only appear in unstressed syllables (Kkese & Petinou, 2017). More specifically, while stressed vowels tend to retain their quality and length, unstressed vowels are shortened and reduced to a schwa (Katamba, 2009). Furthermore, English vowels are shorter before voiceless consonants than before voiced consonants, without a change in their quality (Lengeris, 2009a). Lengeris (2009a) confirmed this, as he found that consonantal context strongly affected the duration of the English vowels, but not their quality, as the F1 and F2 values remained very similar across contexts. Importantly, there is great variation both within and across the native varieties of any language, meaning that even native speakers of a language do not necessarily speak in the same way (Baese-Berk, McLaughlin, & McGowan, 2020).

In addition to phonemic inventory differences, the two languages also differ in orthographic patterns. The following section describes possible orthographic effects arising from these differences.

2.4 The Role of Orthography

Orthographic differences between English on the one hand and SMG and CYG on the other should also be taken into consideration when examining L2 pronunciation by these learners, since orthography can influence the way a word is perceived and produced (Bassetti, Escudero, & Hayes-Harb, 2015; Nimz & Khattab, 2020). The SMG orthography is transparent, as it has grapheme-to-phoneme correspondence, meaning that when seeing a Greek word, it is apparent how to pronounce it (Coutsougera, 2007; Koutsoudas & Koutsoudas, 1962). Even though there are violations to this one-letter-to-one-phoneme correspondence as SMG has a surplus of letters for vowels, each letter or digraph only has one reading (Coutsougera, 2007). In addition, there are generally specific and straightforward rules to address any discrepancies between letters and sounds, meaning that a written form can only be read in one way (Coutsougera, 2007). It should be noted that CYG does not have an established orthography, although based on the SMG orthography, some conventions were developed in CYG written texts for representing sounds that exist in CYG but not in SMG, reflecting aspects of the CYG phonetics and phonology (Arvaniti, 1999a, 2006; Simaki et al., 2015). Importantly, the grapheme-to-phoneme correspondence is maintained in these cases as well.

On the other hand, English has a non-transparent orthography, as one letter can be used to represent more than one sounds, and one sound can represent more than one letters (Coutsougera, 2007). As opposed to SMG, the English orthographic system has fewer letters than corresponding sounds (Coutsougera, 2007). It should be noted that there are rules to account for some discrepancies between letters and sounds in English as well, but not all discrepancies are accounted for by these rules, meaning that pronunciation cannot be reliably predicted from the spelling and vice versa (Coutsougera, 2007; Wells, 2005). This may lead non-native and native speakers alike to make inappropriate inferences from spelling, causing them to mispronounce English words (Coutsougera, 2007; Wells, 2005).

According to Koda (1989), there is a strong relationship between orthography and cognition, and the strategies of phonological coding used in the processing of the L1 are transferred to the processing of the L2, meaning that SMG and CYG learners of English are

likely to simply transfer their L1 strategies to the L2. Furthermore, Bassetti (2006) argued that the mental representations of L2 phonology in beginner learners is affected by the L2 orthography. Therefore, being used to the association of grapheme to phoneme, SMG and CYG learners may try to associate a Greek phoneme with graphemes that have the same form in both Greek and English or to find an association between phonemes and graphemes, despite knowing that this does not apply to English (Koutsoudas & Koutsoudas, 1962).

A number of studies have already investigated orthographic effects in L1 or L2 English and found that orthography affects the production of English segments by native speakers of languages with transparent orthographies. For example, Hayes-Harb, Nicol, and Barker (2010), Rastle, McCormick, Bayliss, and Davis (2011), and Bassetti (2017) found that orthographic forms can interfere with L2 speech perception and production, even when learners are not presented with the orthographic form. Orthographic effects are also reported in Stoehr and Martin (2022), who found that presenting L2 vowel phonemes alongside L1-incongruent orthographic forms may have detrimental effects on both the production and perception of isolated speech sounds. The results of Erdener and Burnham (2005) also suggest that learners whose L1 has a transparent orthographic system are influenced to a larger extent by the orthographic systems of the L2, i.e. they are more likely to be misled by orthography if it does not correspond to the L2 phonological system in a straightforward way. This is in line with other perception studies as well (e.g. Dornbusch, 2012; Simon, Chambless, & Alves, 2010).

At the same time, some studies report no orthographic effects on L2 sound acquisition. For instance, Immonen, Peltola, Tamminen, Alku, and Peltola (2023) examined the effect of orthography on Finnish children's production of the L2 Swedish vowel /ʉ/ and found that these younger learners were not misled by orthographic cues. The researchers note that the fact that orthographic cues did not hinder L2 production learning in this case was contrary to previous findings including Finnish adults (Peltola, Tamminen, Alku, & Peltola, 2015), who were found to rely on the orthographic rather than the acoustic cue, i.e. they produced what they read instead of what they heard, thereby showing sensitivity to the orthographic cues.

Having reviewed the vowel systems of SMG, CYG and English as well as possible effects of orthography on the perception and production of L2 vowels, the following sections turn to the influence of the L1 and the processes involved in L2 segmental perception and production

further, within the framework of the Speech Learning Model (SLM) and the Perceptual Assimilation Model (PAM).

2.5 L2 Production and Perception

The fact that speakers from the same L1 background often produce L2 speech that is distinctive in terms of segmental and prosodic properties suggests that many aspects of their speech are directly related to the influence of their L1 (Derwing & Munro, 2015). Indeed, the importance of the link between the L1 and the L2 sound inventories as a factor that may have a role to play in the difficulties faced by L2 learners is widely recognised (e.g. Aliaga-Garcia & Mora, 2009; Alispahic, Mulak, & Escudero, 2017; Lengeris & Hazan, 2007, 2010; Lengeris, 2018, among others).

This is reflected even in the early stages of research on L2 phonological acquisition with the emergence of the Contrastive Analysis approach during the 1950s-1960s, which included the systematic comparison of the inventories of the L1 and the L2 in order to predict problematic areas for learners as concerns L2 segments (Hammerly, 1982; Kkese & Petinou, 2017; Lado, 1957; Lehn & Slager, 1959; Stockwell & Bowen, 1965). One limitation of this approach was that it could not make accurate predictions as to the acquisition of L2 segments that are not found in the L1 inventory; even though it argued that such sounds would be problematic, L2 learners were in fact found to approach native-like attainment of these sounds in some cases (Kkese & Petinou, 2017). Koutsoudas and Koutsoudas (1962) recognised this problem and argued that learners will find it more difficult to avoid using a native instead of an L2 sound if there is a close association between them. In other words, Koutsoudas and Koutsoudas (1962) argued that if the L1 and L2 phonemes are similar, the L2 sounds will be more difficult to master, as replacing or breaking old habits is more difficult than learning a new set of language habits.

This is also reflected in current models of L2 speech perception and production, according to which the relationship between the L1 and L2 sound inventories may enable predictions as to whether an L2 sound will be difficult for learners. The SLM (Flege, 1995, 2002) and its recent revision, SLM-r (Flege & Bohn, 2021), and the PAM (Best, 1995) and PAM-L2 (Best & Tyler, 2007) are the dominant models accounting for the formation of new categories by learners, in an attempt to predict difficulties in L2 segmental learning and explain how the L1 and L2 phonetic subsystems of learners interact with each other, facilitating or

inhibiting L2 learning. Both of these models (and their variations) support that the L1 has a significant influence over L2 phonological acquisition, i.e. that learners tend to make segmental errors that are L1-specific (e.g. Flege et al., 2006; Jia, Strange, Wu, Collado, & Guan, 2006; Qian et al., 2018). Therefore, the two models also support that difficulties in the perception and production of L2 segments are at least to some extent predictable from the L1, and more specifically from the acoustic similarity or dissimilarity between the L1 and L2 phonemes.

Furthermore, both models agree that the learning abilities used by children in learning an L1 or L2 remain available to adults learning an L2, and share the aim of investigating and uncovering the way in which L2 or unfamiliar phonetic contrasts are perceived by adults (Best & Tyler, 2007). More specifically, according to these models, the differences between early and late L2 acquisition arise not because of biological differences between children and adults, but because of the fact that adult learners have already established their L1, which in turn influences L2 perception and production (Tyler, 2019). In other words, adult learners tend to perceive the sounds of the L2 in terms of their L1 categories, meaning that L2 perception is biased and shaped by the phonetic system of the L1 (Cenoz & Lecumberri, 1999; Sakai & Moorman, 2018; Tyler, 2019). The main difference between the two models is that the SLM is primarily concerned with production although it incorporates perceptual processes, while the PAM focuses on perception (Georgiou, 2019; Tyler, 2019).

The following sections present the main assumptions and hypotheses of the SLM and the SLM-r (2.5.1), and the PAM and the PAM-L2 (2.5.2). It should be noted that although the models described below do not address EFL learners directly, they can both be applied to the EFL context. However, the fact that students in an EFL context vary in the amount of prior experience they have in the L2 and in the native-speaker input they previously received should be taken into consideration (Tyler, 2019). Section 2.5.3 presents the perceptual assimilation patterns of English vowels to SMG and CYG categories, as reported in previous studies, and 2.5.4 focuses on the weighting of acoustic cues by L2 learners.

2.5.1 Speech Learning Model (SLM) and Revised SLM (SLM-r)

As opposed to the proponents of the Critical Period hypothesis (e.g. DeKeyser, 2000; Granena & Long, 2013; Johnson & Newport, 1989; Krashen, 1975; Long, 1990, 2005; Oyama, 1976; Patkowski, 1990; Payne, 1980; Scovel, 2011; Seliger, 1978), who argue that older learners are

unable to achieve native-like knowledge of L2 phonemes due to maturational constraints, the SLM (detailed in Flege, 1995) and its revised version, the SLM-r (detailed in Flege & Bohn, 2021), support that it is possible for adult learners to develop new phonetic categories for the sounds of the L2, since the mechanisms involved in L1 learning remain intact and accessible throughout L2 speech learning. This is supported by two observations: first, that learners with different Ages of Arrival (AOA) in L2-dominant countries were found to have foreign accent differences despite all of them having an AOA over the suggested critical period (e.g. Flege & MacKay, 2011); second, that some young children learning an L2 maintain a foreign accent even after decades of L2 use (e.g. Flege et al., 2006). For example, Flege et al. (1997a) examined Italian adults with an AOA of 6 years and found that not only were these learners foreign-accented, but the strength of foreign accent in their speech was correlated with frequency of continued L1 use.

According to Flege (1995, 2002, 2007, 2018) the L1 and L2 phonetic categories exist in a common phonological space in which they interact through the processes of phonetic category assimilation and phonetic category dissimilation (Flege, 2002, 2003, 2007; Flege, Schirru, & MacKay, 2003). Phonetic category assimilation occurs when the creation of a new category is blocked due to equivalence classification, i.e. the perception of an L2 sound as phonetically similar to an L1 sound. It is usually found at the early stages of L2 learning, and results in the production of L2 sounds using the same articulatory routines as those for the production of L1 sounds (Flege et al., 2003; Flege, 2002, 2007, 2008). Phonetic category dissimilation occurs when a new phonetic category is ultimately formed for an L2 sound, in which case the phonetic space becomes more crowded and the phonetic categories of the L1 and the L2 disperse so as to maintain contrast between the sounds of the two languages (Flege et al., 2003; Flege, 2002, 2003, 2007). The assumption that new category formation is possible for L2 learners irrespective of age of first exposure to the L2 is reiterated in the SLM-r, although it is stressed that L2 learners cannot match monolingual L1 speakers of the target L2 because of two main reasons: firstly, because the two phonetic subsystems of bilinguals exist in a common phonological space and therefore inevitably interact, and secondly, because the input an L2 learner receives is inevitably different to the input that native speakers receive (Flege & Bohn, 2021).

According to the SLM, the formation of a new category for an L2 sound depends on two factors. The first is the degree of development of L1 categories as age increases, with more developed L1 categories becoming more likely to subsume L2 sounds, blocking the establishment of new phonetic categories (Flege, 1995, 2007). However, it is assumed that even late learners can modify their productions, given that phonetic learning remains possible throughout the lifespan (Flege, 1995, 2007). The second is the perceived phonetic similarity/dissimilarity between an L2 sound and the closest L1 sound, in which case the establishment of a new category is more likely when the L2 sound is perceived to be more distant from the closest L1 sound (Flege, 1995, 2007). Therefore, the likelihood of new category formation decreases as the L1 categories become more robust with age, which makes it more difficult for leaners to discern phonetic differences between similar L2 sounds and the closest L1 sound (Flege, 1995, 2007).

The SLM-r revises this proposition and suggests, instead, that new category formation depends on three rather than two factors, i.e. the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, which remains unchanged from the SLM, the quantity and quality of L2 input, and the degree of L1 category precision of the closest L1 category at the time when L2 learning begins. According to the newly introduced "L1 category precision" hypothesis, individuals who have relatively precise L1 phonetic categories will be able to perceive phonetic differences between an L2 and an L1 sound better, compared to learners that have relatively imprecise L1 categories, thereby increasing the likelihood of new category formation. Precision is defined as having little within-category variability and relatively large between-category distances, and it is argued to lead to finer discrimination abilities. According to Flege and Bohn (2021, p. 65), "L1 category precision generally increases through childhood and into early adolescence, but important individual differences exist at all ages". This hypothesis can also account for the individual differences found among learners.

In terms of perceptual processes, in the initial formulation of the model, Flege (1995, 2003) argued that perception and production are linked in a way that production performance is inhibited by insufficient perceptual ability (see also Baker & Trofimovich, 2001; Detey & Racine, 2015). The SLM hypothesised that upon discerning the subtle phonetic differences between similar L1 and L2 sounds, learners may be able to form a new phonetic category for L2 sounds, which can facilitate, but not necessarily guarantee, accurate L2 production (Flege,

1995). Therefore, non-native sounds that are more dissimilar to the closest L1 sound are expected to be easier to perceive as different, meaning that learners are more likely to establish separate categories from the existing L1 categories for these sounds, provided that they receive adequate input. Furthermore, the SLM hypothesised that speech perception can be improved given sufficient exposure, in which case production will improve in accuracy over time as well (Flege, 1995). Importantly though, segmental perception is not always perfectly aligned with segmental production; for instance, learners may not update their production even when accurate perception of a sound is achieved, since, in addition to perception, motor programs are also required for successful production (Flege, 1999; Leather & James, 1991).

The nature of the interaction between perception and production has been the subject of investigation of various previous studies, although a consensus has yet to be reached (Melnik-Leroy, Turnbull, & Peperkamp, 2022). For example, many studies report at least a modest relationship between the two modalities (e.g. Baker & Trofimovich, 2006; Bettoni-Techio, Rauber, & Koerich, 2007; Jia et al., 2006; Kluge, Rauber, Reis, & Bion, 2007; Melnik-Leroy et al., 2022; Zhang & Peng, 2017), while others found partial (e.g. Levy & Law, 2010) or no correlation between them at all (e.g. Kartushina & Frauenfelder, 2014; Peperkamp & Bouchon, 2011). In addition, the direction of the link is not yet established. The initial hypothesis of the SLM that accurate perception precedes accurate production is supported by some experimental studies (e.g. Casillas, 2019; Flege, 1993; Melnik-Leroy et al., 2022; Nagle, 2018), which found that perceptual improvements preceded production improvements indicating that production lags behind perception. However, other studies have found accurate production despite an inaccurate perception of an L2 contrast (e.g. Bohn & Flege, 1997) or mixed results such as better perception for some sounds and better production for others (e.g. Hao & de Jong, 2016). According to Melnik-Leroy et al. (2022), the inconsistency in the findings may be due to various factors, including methodological decisions, type of target L2 sounds, the relationship between target L2 sounds and L1 sounds, the level of proficiency of learners or individual differences among participants.

Such findings prompted one important revision in the SLM-r concerning the relationship between perception and production. More specifically, while the SLM assumed a unidirectional relationship between the two modalities, with perception shaping accuracy in production, the SLM-r has revisited this relationship, arguing for a bidirectional link, whereby

"segmental production and perception coevolve without precedence" (Flege & Bohn, 2021, p. 64).

Finally, a crucial observation incorporated in the SLM-r is that, despite the implicit assumption that all native speakers of a language share identical or at least very similar phonetic categories, individual differences in the perception and production of L1 phonetic categories also exist depending on the input distributions they received during L1 development, or on different degrees of precision in how L1 categories are defined. These individual differences in L1 phonetic categories may have an effect on individual learners' L2 speech learning (Flege & Bohn, 2021).

2.5.2 The Perceptual Assimilation Model (PAM) and the PAM-L2

The PAM (Best, 1995) and the PAM-L2 (Best & Tyler, 2007) also investigate L1 interference and the ways in which non-native phonemes are perceived by adult learners, providing category assimilation patterns across languages which can enable predictions as concerns the degree of L2 discrimination accuracy. The PAM was initially developed to explain the patterns of assimilation of non-native sounds to L1 categories by naïve listeners. This was later applied to L2 learning (PAM-L2), in an attempt to predict the degree of difficulty an L2 contrast might pose for learners based on the perceptual assimilation patterns (Best & Tyler, 2007). The PAM-L2 assumes a shared perceptual system for all languages of a learner and the assimilation patterns of L2 phonemes onto L1 categories determine the success of a learner in detecting new phonological contrasts (Tyler, 2019). According to the researchers, when a new sound is encountered, learners attempt to assimilate it to perceptual categories already in place for the L1. This means that new sounds are perceived through a "perceptual sieve" based on L1 knowledge. If an L2 phoneme is perceived to be very similar to an L1 phoneme, then it is more likely that it will be assimilated to the L1 category (Best & Tyler, 2007; Best, 1995). Importantly, L2 contrasts pose differing degrees of difficulty for L2 learners based on the similarities and differences between the L2 and L1 contrasts.

According to the PAM, there are six patterns of assimilation of non-native phonological contrasts onto native categories by naïve listeners, which can also be extended to L2 perceptual learning, and can enable predictions as to how accurately the contrasts will be discriminated. These are Two-Category Assimilation (TC type), Category-Goodness Difference (CG type), Single-Category Assimilation (SC type), Uncategorised-Uncategorised (UU type),

Uncategorised-Categorised (UC type), and Non-assimilated (NA type) (detailed in Best & Tyler, 2007 and Best, 1995).

According to the models, TC contrasts are expected to be easier for learners to discriminate, followed by CG contrasts, where L2 phonemes are assimilated to a single L1 phonological category but with a different goodness of fit, which increases the likelihood for learners to acquire a new phonological category (Best & Tyler, 2007; Best, 1995). On the other hand, SC assimilations, which involve two sounds being assimilated to a single L1 category as equally good or poor exemplars of it, are predicted to be particularly challenging for learners, even if they are exposed to high-quality native speaker input (Tyler, 2019). This difficulty often arises when the L1 phonological category permits phonetic variability to the extent that it encompasses the phonological contrast in the L2, i.e. when allophones in the L1 are phonemes in the L2, in which case the L2 phonemes are difficult to perceive as two distinct phonemes (Tyler, 2019). Such sounds may need targeted training so that learners can detect the differences between them (Tyler, 2019). The discrimination of UC and NA contrasts is also expected to be very good, whereas the discrimination of UU contrasts, in which neither of the L2 sounds falls within an L1 category, may vary from poor to very good, depending on their similarity to an L1 category (Best, 1995; Tyler, 2019).

The predictions of the PAM and PAM-L2 have been supported by various previous studies. For instance, Tyler, Best, Faber, and Levitt (2014) examined the perceptual abilities of American listeners in three unfamiliar languages and found that the contrasts falling into the TC assimilation type were better discriminated than contrasts of the CG type, which were in turn better discriminated than contrasts of the SC type, supporting the assumptions of the PAM. In addition, Lengeris (2009a, 2009b) examined whether discrimination accuracy could be predicted from perceptual assimilation patterns according to PAM categories and found that TC contrasts were not problematic for his learners, who showed some difficulty with UC and CG contrasts and had the most difficulty in discriminating SC contrasts. The predictions of the PAM are supported by Georgiou (2019) as well, who examined CYG learners of English and found the assimilation patterns to be mostly compatible with their discrimination accuracy as proposed by the PAM, i.e. poor discrimination was found in SC assimilation types, moderate discrimination was found in CG patterns and moderate to good discrimination was found in the UC type.

Since the PAM-L2 makes hypotheses for learners in immersion contexts where the L2 is dominant, Tyler (2019) sought to describe how the principles of the model can be applied to learners in an FLA context. According to Tyler (2019), the predictions of the PAM-L2 for the acquisition of TC contrasts remain the same for FLA contexts. However, CG assimilations will be less likely to be acquired than in an immersion setting, especially if the foreign-accented input received does not differentiate the L2 phonemes. Finally, SC assimilations, which are already difficult to acquire in immersion situations, will be even more difficult to acquire in the EFL classroom (Tyler, 2019). These predictions are supported by Balas (2018), who examined the perception of English vowels by Polish EFL learners and found that discrimination rates depended on the assimilation types.

Finally, it should be noted that even though there is an assumption of a link between perception and production in PAM and PAM-L2, none of them directly discusses the link between the two modalities, nor were they developed to explain how the link between the two modalities develops and evolves or how they interact throughout the L2 learning process (Nagle & Baese-Berk, 2022).

Having presented the two current models of L2 speech perception and production, the following section will focus on the perceptual assimilation patterns of L2 English vowel contrasts to L1 SMG and CYG categories, as detailed in Lengeris (2009a) and Georgiou (2019), respectively.

2.5.3 Perceptual Assimilation Patterns of English Vowel Contrasts by Greek Learners

Lengeris (2009a) investigated SMG learners' perceptual assimilation and discrimination of English vowels, exploring for the first time the perceived relationship of the vowels in the two languages. The researcher found that SMG learners used both durational and spectral cues in assimilating L2 vowels to L1 categories, despite the fact that durational cues are not used contrastively in SMG. The results of this study show that two or more English vowels are assimilated to the same SMG category, albeit with varying degrees of fit, i.e. some English vowels were assimilated to one SMG category consistently and they were judged to be good exemplars of it, and others were judged to be poor exemplars of an SMG category, sometimes heard as falling between two SMG categories. The assimilation patterns observed were as follows: English /iː/ and /ɪ/ were assimilated to SMG /i/, English /e/ and /ɜː/ were assimilated

to SMG /e/, English /æ/ and /ʌ/ were assimilated to SMG /a/, English /ɑ:/, /ɒ/ and /ɔ:/ were assimilated to SMG /o/ and English /ʊ/ and /u:/ were assimilated to SMG /u/.

Georgiou (2019) conducted a similar study examining the perceptual assimilation patterns and goodness of fit ratings of English vowels by child native speakers of CYG. The findings of this study also show that CYG learners assimilate more than one vowel of the L2 to one category in their L1, as a result of the fact that their L1 vowel inventory is smaller than the L2 vowel inventory, since the main assimilation types occurring were either CG or SC which indicate assimilation of two L2 contrasting vowels to a single L1 phonological category. This is consistent with previous studies investigating the perception of vowels by learners whose L1 involves a smaller vowel inventory than their L2 (e.g. Escudero, 2005 for Spanish learners; Lengeris, 2009a, 2009b for SMG learners).

Although goodness of fit ratings varied in this study as well from poor to very good, the assimilation of English vowel contrasts to CYG categories was found to be as follows: English /ɪ/ and /i:/ were assimilated to CYG /i/, English /e/ and /ɜ:/ were assimilated to CYG /e/, English /æ/, /ʌ/ and /ɑ:/ were assimilated to CYG /a/, English /ɒ/ and /ɔ:/ were assimilated to CYG /o/ and English /ʊ/ and /u:/ were assimilated to CYG /u/. Therefore, it is evident that the English vowel contrasts were assigned to a single L1 phonological category, which according to the PAM means an SC (when the two vowels are equally good or equally poor exemplars of the L1 category) or a CG assimilation type (when one of the vowels of the contrast is a good exemplar and the other is a poorer exemplar of the L1 category) (Georgiou, 2019).

A comparison of the above findings shows that the assimilation patterns of English vowel contrasts to the SMG and CYG phonological categories are very similar, i.e. the vowels /i:, ɪ/, /e, ɜ:/, /æ, ʌ/, /ɒ, ɔ:/ and /ʊ, u:/ were mostly assimilated to Greek /i/, /e/, /a/, /o/ and /u/ respectively. One difference is that in Georgiou (2019), the /ɑ:/ vowel was assimilated to the Greek category for /a/ instead of /o/ as Lengeris (2009a) found. This may be attributed to the fact that participants in Lengeris (2009a) were adults with more experience in English (10-15 years), while in Georgiou (2019), participants were young learners with 1-5 years of learning experience. A second explanation could be the difference in the variety spoken by the participants and the acoustic differences between the vowels in the two varieties which may have influenced the way in which the native speakers of each variety classify the English vowels.

Similar results are reported for Catalan and Spanish learners of English in various previous studies (e.g. Aliaga-Garcia & Mora, 2009; Carlet & Cebrian, 2014; Cebrian, 2019; Cebrian, Carlet, Gorba, & Gavaldà, 2019); the lack of a tense-lax distinction in the L1 of learners was found to cause the assimilation of the English vowels to a single L1 category, as they occupy an area of the vowel space that is filled by a single L1 category. These studies have shown that English /iː/ and /ɪ/ tend to be assimilated to Spanish /i/, English /ɛ/ and /ɜː/ tend to be assimilated to Spanish /e/, and English /ɑː/, /ʌ/ and /æ/ to Spanish /a/ (Aliaga-Garcia & Mora, 2009; Carlet & Cebrian, 2014; Cebrian et al., 2019; Cebrian, 2019).

Importantly, the perceptual patterns reported in Georgiou (2019) can enable predictions as to the production of English vowels by CYG learners as well, since these learners are expected to use their five L1 vowels for the production of the L2 vowels similarly to their perceptual patterns. Georgiou (2019) reports a CG or UC assimilation pattern for the English vowel contrasts /ɪ/-/iː/, /e/-/ɜː/ and /ʊ/-/uː/, and an SC assimilation pattern for the English vowel contrasts /æ/-/ʌ/ and /ɒ/-/ɔː/. Assuming that better discrimination will lead to better production, and based on the PAM's prediction that SC contrasts are the most difficult to discriminate, it is expected that CYG learners of English will find it more difficult to accurately produce the vowels /æ/-/ʌ/ and /ɒ/-/ɔː/.

It should also be noted, however, that L2 vowel perception can be strongly affected by the place of articulation of neighbouring consonants, promoting different assimilation patterns in the vowels, although this does not affect all vowels equally. Bohn and Steinlen (2003), for example, examined the identification of the 11 Standard British English monophthongs by Danish listeners in 3 different contexts (/hVt/, /dVt/ and /gVt/) and found that the perceptual assimilation of /ɪ ɛ ʊ ʌ/ was strongly affected by consonantal context, while /ɔː ɜː/ were not much affected by context and /iː uː æ ɒ ɑː/ were very consistently identified across contexts. In addition, coarticulation effects may vary depending on the L1; for instance, Dutch vowels are minimally affected by flanking consonants, and therefore, Dutch learners of English are expected to show greater difficulties when encountering contextual variation (Bohn & Steinlen, 2003).

At the same time, such effects are expected to occur at the initial stages of acquisition, since more experienced learners were found to disregard contextual variation and perceive L2 vowels more consistently (e.g. Balas, 2018; Levy & Strange, 2008). For example, Levy and

Strange (2008) also found consonantal effects in the perceptual discrimination of L2 French vowels by L1 American English listeners, in that they found that vowels in a bilabial context (/bVp/) were more accurately discriminated than vowels in an alveolar context (/dVt/). However, these effects were found in inexperienced learners only, while no context effects were found in more experienced learners. Furthermore, Balas (2018) examined the perception of L2 English vowels by advanced Polish FLA learners in 3 consonantal contexts (bilabial, alveolar and velar) in order to assess the effects of place of articulation of adjacent consonants on vowel discrimination and categorisation, and contrary to previous studies with lower-proficiency learners, the researcher found limited effects of context in the perception of the target English vowels, i.e. the place of articulation of the neighbouring consonants did not seem to play a crucial role in the perception of English vowels. The researcher hypothesised that this result was due to the fact that the participants were advanced learners, who had possibly mastered L2-specific coarticulatory variations (Balas, 2018).

The following section will focus on how the L1 can influence acoustic cue-weighting in the perception of L2 segments.

2.5.4 Weighting of Acoustic Cues

While some aspects of non-native speech are common across any combination of L1 and target L2 (e.g. non-native speech is typically slower than native speech), many aspects of it largely depend on the L1 background of the speaker and the target L2, as reflected in most L2 speech acquisition models (Baese-Berk et al., 2020). Establishing a new phonetic category involves specifying how multiple acoustic cues are integrated and weighted; the value of these cues may also vary as an effect of context, stress and speaking rate, among other factors (Flege & Wayland, 2019). Since different languages employ acoustic cues differently to signal phonological contrast, the weighting of these acoustic cues in the perception of L2 learners differs as an effect of the L1, posing additional difficulty for L2 learners (Holliday, 2010).

One explanation for this difficulty was provided by the Feature Hypothesis (proposed by McAllister, Flege, & Piske, 2002), according to which L2 learners only have access to cues that are used contrastively in their L1, making it more difficult to form a new category based on a phonetic cue that is not used in the L1. This was supported by McAllister et al.'s (2002) results, as their participants' success in acquiring Swedish quantity distinctions was related to the importance of durational cues in their respective L1s, suggesting that the difficulty in acquiring

an L2 contrastive feature lay on the relative importance of that feature in the L1. Based on these results, the Feature Hypothesis was initially supported by the SLM, but was then abandoned by the SLM-r, which adopts a "full access" hypothesis instead, supporting the view that L2 learners can access features that are not exploited in their L1 (Flege & Bohn, 2021). More specifically, according to the SLM-r, as long as an L2 sound remains perceptually linked to an L1 category, L1 cue weighting patterns will remain stronger. However, in newly developed L2 phonetic categories, the cue weighting patterns are expected to develop in the same manner as in L1 acquisition, i.e. they are expected to be based on the reliability of cues when it comes to rapid and accurate categorisation of phonetic segments (Flege & Bohn, 2021).

This is in line with Holliday (2010), who suggested that since the same basic auditory function is available to all normal-hearing language users, it is more likely that L2 learners are merely not used to attending to some acoustic cues that are not employed in their L1, rather than that these cues are physically unavailable to them. For example, the Japanese learners in Yazawa, Whang, Kondo, and Escudero (2020) were able to shift their cue weighting patterns and increased their reliance on spectral cues when they thought the stimuli were English vowels, even though their L1 only employs durational cues. Therefore, in order to reliably discriminate the members of an L2 contrast, L2 learners must learn to attend to the relevant acoustic cues (Holliday, 2010; Kkese & Petinou, 2017).

Importantly, previous research has shown that L2 learners are sensitive to durational cues, even when their L1 does not use the duration feature contrastively (e.g. Cebrian, 2006; Escudero & Boersma, 2004; Escudero, Benders, & Lipski, 2009; Lengeris, 2009a, 2009b; Morrison, 2002). For example, whereas native English speakers rely primarily on spectral cues in identifying their L1 vowels (Cebrian, 2006; Hillenbrand et al., 2000; Zhi & Li, 2021), Catalan and Spanish learners were previously found to over-rely on durational cues in perceiving and producing L2 English vowels, even though vowel duration is not used contrastively in their L1 (Aliaga-Garcia & Mora, 2009; Cebrian, 2006, 2007; Kondaurova & Francis, 2010; Mora & Fullana, 2007, among others). The same has been observed in studies with other groups of learners whose L1 has fewer monophthongs than the L2 and does not contrast tense and lax vowels, such as Polish (Bogacka, 2004; Rojczyk, 2010) and Chinese Mandarin (Zhi & Li, 2021) learners of English. One explanation for this is offered by Bohn's (1995) Desensitisation

Hypothesis, according to which learners tend to rely on durational cues when they cannot differentiate L2 vowels using spectral cues, irrespective of whether durational cues are used contrastively in their L1. According to Bohn (1995), this happens because duration is a language-independent, salient cue that is easily accessible.

Sensitivity to durational information in the L2 has also been reported for SMG learners of English (e.g. Lengeris & Hazan, 2007), even though their L1 does not use this feature. Lengeris (2009a, 2009b) also found that SMG learners have access to durational cues, although his results suggest that they do not merely rely on them, but rather attend to both spectral and durational cues when perceiving English vowels. Similarly, the CYG young learners in Georgiou (2019) were found to use both spectral and durational features in categorising the L2 vowels into their L1 categories. More specifically, even though spectral cues are not used in the L1 either, his results suggest that these learners primarily relied on spectral similarities between the L1 and L2 vowels, turning to durational features only in the absence of spectral cues.

Having discussed the influence of the L1 and the current models of L2 speech perception and production, the following sections turn to other individual factors that have been argued to play a role in L2 phonetic acquisition and degree of foreign-accentedness in L2 speech, focusing specifically on language use patterns, motivation and input.

2.6 Individual Differences in the Perception and Production of Non-Native Sounds

The age at which L2 acquisition starts is considered to be a determining factor in L2 phonological attainment. More specifically, it has been argued that as the AOL increases, the acuity to perceive the L2 phonetic categories declines (Best & McRoberts, 2003; Scovel, 1969). In addition, it has been suggested that the development of the L1 speech sound inventory inhibits the acquisition of unfamiliar sounds, making the perception and production of foreign speech sounds more difficult after the first six months of life of an infant (e.g. Best & McRoberts, 2003).

Although the premise “the earlier, the better” when it comes to L2 learning is widely accepted, previous research shows that an early AOL does not automatically lead to native-like L2 speech (Flege et al., 1997a) but also that even learners with an AOL higher than 12 years can still achieve learning to speak the L2 without a foreign accent (Flege et al., 1995).

Therefore, as opposed to the assumption that there is a biologically determined Critical Period for language learning (Johnson & Newport, 1989; Lenneberg, 1967; Schouten, 2009), meaning that maturational constraints and a loss of neural plasticity of the brain prevent learners from achieving a native-like accent in an L2 after childhood or adolescence (Granena & Long, 2013; Long, 2005; Oyama, 1976; Payne, 1980; Scovel, 2011), some researchers (e.g. Flege, 1999, 2008; Moyer, 1999, 2004; Singleton, 2003, 2005) suggest examining other factors to better understand L2 speech acquisition, since the age of the learners is usually confounded with other factors that may have an effect on L2 pronunciation. These include both extrinsic factors (e.g. AOL, LOR in an L2-speaking community, length of formal instruction, quantity and quality of L2 input, amount of L2 use/L1-L2 use patterns, etc.) and intrinsic factors (e.g. motivation, memory, language learning aptitude, etc.) (Moyer, 2013; Piske et al., 2001; Piske, 2007; Rato & Carlet, 2020).

Some researchers argue that native-like pronunciation can be achieved by late L2 learners as a result of a combination of these factors (e.g. Birdsong & Molis, 2001; Bongaerts et al., 1997; Bongaerts, Mennen, & Slik, 2000), possibly explaining the individual differences found among participants in previous studies. Although various factors have been examined in the literature, some of the most prominent ones are AOL, LOR in an L2-speaking country and amount of continued L1 use (e.g. Piske et al., 2001), motivation (e.g. Moyer, 1999), input (e.g. Flege, 2008), attitude (Elliot, 1995; Moyer, 2007) and instruction (e.g. Krzonowski et al., 2015).

Indeed, the examination of these factors on L2 speech performance in previous studies suggests that L2 speech learning is mainly influenced by four inter-related factors, i.e. AOL, amount of L1-L2 use, quantity and quality of L2 input and formal instruction in the perception and production of L2 sounds (Piske, 2007). For example, Rato and Carlet (2020) investigated the effects of AOL, length of formal instruction, L2 language use and vocabulary size on the perceptual performance of Portuguese learners in L2 English vowel identification. Based on the results, the only variable affecting vowel identification was amount of L2 language use. Further correlation analysis showed that the amount of daily L2 use was positively associated with the identification of some vowels (/ɪ/, /ɛ/ and /æ/), showing that higher L2 use was related with more accurate identification of these target vowels.

Furthermore, some studies (e.g. Birdsong, 2007; Bongaerts, 1999; Bongaerts, Planken, & Schils, 1995) also provide encouraging results for late learners. For instance, the Dutch learners of English and French in Bongaerts et al. (1995) and Bongaerts (1999), respectively, were found to be "highly successful" in the production of sounds that are considered challenging for these learners, leading Bongaerts (1999) to conclude that three coexisting factors are likely to have played a crucial role in this success, i.e. a high degree of motivation, massive exposure to the L2, and intensive training in L2 perception and production skills. Birdsong (2007) also observed that successful learners were highly motivated and had received phonetic training.

Importantly, such results demonstrate that achieving native-like speech is not impossible for late L2 learners who learn the L2 in a classroom setting. The present research is concerned with only some of the factors that may affect L2 acquisition, focusing on those that can directly lead to changes in pronunciation instruction to improve performance, especially for adult learners in an EFL context. These are discussed in the following sections.

2.6.1 Language Use Patterns

In terms of the influence of L1-L2 use patterns on overall degree of foreign accent, previous studies have observed a correlation between high L2 use and milder foreign accents and an inverse correlation between high L1 use and foreign accent (Flege et al., 1997a; Flege et al., 1999a; Flege, 2008; Flege, MacKay, & Meador, 1999b; Piske et al., 2001, among others). The effect of language use patterns was initially examined by Suter (1976) and Purcell and Suter (1980) through an investigation of the self-estimated habits of learners in using their L2 at home and at work or school. Although no significant effect of this factor alone was observed, language use patterns in combination with LOR and cohabitation with native speakers were found to be important in predicting degree of L2 foreign accent. A significant effect of L2 use on the perception and production of three vowel contrasts (/iː-ɪ/ /ɛ-æ/, /ʌ-ɑ/) by Mandarin learners of English is reported in Jia et al. (2006) as well. Furthermore, Guion, Flege, and Loftin (2000) report that L1 use has an effect on L2 production, as they found that the group of participants who used their L1 more, had stronger L2 accents compared to the participants with lower L1 use. Piske et al. (2001) also found that amount of continued L1 use had an effect on the degree of foreign accent in L2 speech.

However, research investigating the effect of L2 use on degree of foreign accent is not conclusive, as other studies (e.g. Elliott, 1995; Thompson, 1991) did not find a significant effect of this factor. The conflicting results in the literature may be due to the fact that this factor is difficult to measure consistently across participants and studies, and it is usually confounded with other factors such as LOR, AOL or differences in L1 background (Piske et al., 2001). Importantly, previous studies have focused on bilingual speakers residing in the L2-speaking community. In EFL settings, L1 use has only been examined in relation to its use in the EFL classroom (e.g. Bozorgian & Fallahpour, 2015; Pan & Pan, 2010). To my knowledge, the only previous study examining the effect of language use patterns on L2 pronunciation in an EFL context is Dimitriou (2019), who investigated L2 English rhotic production by CYG learners grouped according to their language use patterns and quality and quantity of input received during their secondary education in Cyprus, and found an advantage for learners with additional input and L2 use.

2.6.2 Motivation

Turning to the influence of motivation, it seems reasonable to assume that L2 learners who are concerned about their pronunciation are more likely to perform better, at least partly because these learners are more likely to seek out opportunities to use and improve their skills in the L2 (Gilakjani, 2012). At the same time, while some learners may be aware that their speech is foreign-accented, they may not have the motivation to improve their performance as long as their ability to communicate effectively is not compromised (Moyer, 2007).

Previous studies examining the role of motivation report that this factor may have some effect on degree of L2 foreign accent, although it has not been found to be strong enough to "automatically lead to accent-free L2 speech" (Piske et al., 2001, p. 202). For example, Purcell and Suter (1980) and Elliott (1995) report that motivational variables, i.e. learners' concern about L2 pronunciation, can play a significant role on degree of foreign accent. Strength of concern about pronunciation was also found to be a variable strongly correlated with pronunciation accuracy in Suter (1976), while Bongaerts et al. (1997) report that five out of the eleven participants in their study, who were highly motivated late learners of English, reached pronunciation ratings comparable to native speakers' ratings. Yousofi and Naderifarjad (2015) also found that motivation was significantly correlated with pronunciation

skill as measured through six tasks (wordlist reading, sentence reading, dialogue reading, text reading, picture description, interview and role playing) and rated by three professional raters.

Attitude towards acquiring a good or native-like pronunciation was also found to be an important variable positively correlated with pronunciation accuracy (Elliot, 1995). For example, Moyer (2007) investigated language attitudes and their role in degree of foreign accent, and concluded that this is a significant factor influencing pronunciation and that a desire to improve accent, a positive orientation towards the L2 and the L2 culture, comfort with assimilation, and having the intention to reside long-term in an L2-speaking country can all be important factors in improving pronunciation. Furthermore, according to Gilakjani (2012) personal and professional motivation to learn English can have a positive effect on the desire of learners to reach native-like pronunciation.

However, the results of previous studies concerning the effect of this factor are not conclusive. For example, Moyer (1999) examined this variable on late learners of German with high professional motivation, and while there was a strong correlation between professional motivation and foreign accent, none of the participants reached pronunciation ratings similar to those for native speakers. In addition, no significant effect of motivation on degree of foreign accent was found in Thompson (1991).

The inconclusive results of previous studies may be explained by the fact that, similarly to language use patterns, this factor is difficult to measure precisely and consistently across studies, especially since motivation can take many forms. For example, motivation can be instrumental or integrative, i.e. it can refer to linguistic achievement or learning about the culture of the target language, respectively (Gardner & Lambert, 1972; Gilakjani, 2012), and it can be related to an individual's profession or language attitudes, as shown in the studies above. In addition, the lack of consensus may be due to the fact that most studies rely on participants' self-ratings concerning, for example, how important they consider a good pronunciation to be for their work and for their social life, as well as their attitudes towards the L2, which makes it difficult to know the accuracy of the responses (Piske et al., 2001). Importantly, the different results obtained in previous studies and the differences in the success of different groups of learners may also be due to a combination of other factors coexisting with motivation, such as the amount of input that learners were exposed to, the

age at which participants were firstly exposed to the L2, and L1-L2 differences (Piske et al., 2001).

However, the fact that at least a small effect of motivation was found in most studies examining this factor should not be neglected. Further investigations of the effect of motivation can provide insights as to the importance of motivating learners in EFL contexts to aim for native-like proficiency, whilst also acknowledging that motivation alone may not be enough, as Wells (2005) suggests. This is particularly important given that little attention is paid to pronunciation in the EFL classroom (see 2.2).

2.6.3 Input

A widely-researched factor that may have a role to play in L2 production is the amount of native-speaker as opposed to foreign-accented input to which a learner is exposed (Flege & Liu, 2001; Flege & MacKay, 2004; Flege et al., 2003; Flege, 2007, 2008). According to Flege (2008), L2 input differs from L1 input, and it is more variable, as it is usually foreign-accented, especially during the early stages of learning. Therefore, both quantity and quality of input should be taken into consideration when investigating the effect of this factor. According to Flege and Wayland (2019, p. 25) "phonetic categories are *perceptual* representations that speaker-hearers establish over time in long-term memory based on the phonetic input they receive. A phonetic category is defined by *all* of the tokens encountered on the phonetic surface of meaningful speech that have been identified as instances of the category". Therefore, speaker-hearers of a language produce segments with the acoustic properties they have in the input to which they have been exposed, and exposure to different input distributions is likely to lead to individual differences in the phonetic categories (Flege & Wayland, 2019).

The importance of input for L2 speech learning is controversial. While some researchers (e.g. DeKeyser & Larson-Hall, 2005; DeKeyser, 2000) support that input plays a very limited role in L2 pronunciation, being far less important than age of first exposure, others (e.g. Flege, 2008, 2018) argue that the age effects observed in learners with a later AOL are a result of input differences rather than a loss of their capacity for speech learning, assigning much more importance to the quantity and quality of input in successful L2 speech learning. More specifically, Flege (2018) argues that age effects are, in reality, input effects, since age reflects quantity of input and AOA "conditions [learners'] experience in the L2" (p. 919) both in quality

(formal or informal context of learning, contact with native or non-native speakers, etc.) and quantity (shifts in amount of L1 and L2 use due to social factors).

Despite the considerable attention paid to this factor in the literature, its importance remains uncertain. According to Flege and Liu (2001), L2 speakers may learn to accurately perceive and produce L2 sounds, provided that they receive adequate and high-quality native-like input. Dimitriou (2019) compared learners who had attended Greek-speaking state schools with learners who had attended English-speaking private schools during their secondary education in Cyprus and found that CYG learners who had been exposed to the language and to native-speaker input to a larger extent were more successful in producing the English approximant /ɹ/, demonstrating the importance of the quality and quantity of input in L2 segmental production. Hutchinson and Dmitrieva (2022) examined the perception and production of the French vowels /y/ and /u/ by naïve monolingual American English speakers and the effect of a single session of exposure to French film on their performance and found that this method was successful in improving learners' pronunciation of the /y/, but not the more challenging /u/ vowel.

On the other hand, Cebrian (2006) found that two Catalan groups differing in experience with the L2, i.e. learners with an AOL of 20-45 y.o. in the host country vs. English majors living in Catalonia and exposed to different varieties of English, did not differ significantly in vowel identification. Flege and Wayland (2019) also investigated the role of quantity of L2 input on the perception and production of L2 segments by relatively inexperienced Spanish late learners of English and found that increased input (from 0.2 years to 3.0 years) was not sufficient in improving vowel production and consonant discrimination, while only modestly improving vowel discrimination as years of English input increased.

As Flege and Wayland (2019) suggest, there are two possible explanations for this lack of input effect. The first relates to maturational constraints, i.e. the fact that input received after the Critical Period does not make learners' perception and production of L2 segments more native-like, despite the existence of some successful late learners. According to this explanation, and based on Lenneberg's (1967) initial formulation of the Critical Period Hypothesis, which suggests that learners cannot automatically use input just by being exposed to it and that the success enjoyed by some late learners may be due to "conscious and labored effort" (p. 176), successful late learners may have reached native-like productions only

because of a special aptitude for L2 learning or of strong motivation that led them to work hard to learn the L2 pronunciation. The second explanation relates to the continued influence of the L1, which leads to a non-native-like segmental production and perception in the L2, in combination with an inadequate amount of L2 input received. The researchers argue that if the process of L2 speech learning is the same as L1 learning, learners would need the same amount of native-speaker input as monolingual children do in order to establish new phonetic categories.

Importantly, Flege and Wayland (2019) point out that confidently choosing between the two competing explanations necessitates the identification of appropriate and accurate methods to measure quantity and quality of input. Indeed, measuring input reliably is a challenge; current research relies on the self-estimates of participants using questionnaires to measure learners' overall L2 use, an inadequate approach that does not provide insights into how much of this input is foreign-accented (Flege & Wayland, 2019; Flege, 2008). Furthermore, as Flege (2018) points out, there is an over-reliance on LOR as an indication of amount of input received, which is often an unreliable measure of amount of L2 use or input, or the nature of the L2 input. As a result, it is difficult to know how much input L2 learners actually receive, whether this is adequate, or how much exposure is needed for them to achieve intelligibility or native-likeness (Piske & Young-Scholten, 2008). Therefore, the importance of input for L2 speech learning remains unclear.

Finally, input is also closely associated with formal instruction and training (see 2.2 and 2.8), which also contributes to the controversial findings reported in previous studies. For example, Georgiou (2019) found no differences in the perceptual patterns of two CYG groups differing in proficiency level, and therefore concluded that quantity of input alone is inadequate for the acquisition of L2 sounds. The researcher attributed this lack of a difference between the two groups to the exposure of these learners to mostly foreign-accented speech by non-native language teachers, in combination with the lack of attention to pronunciation instruction in the CYG classroom (see 2.2), pointing to the significance of the quality of input in EFL classrooms.

Having examined some of the factors that may affect L2 phonological acquisition, the next section will address the concepts of intelligibility and comprehensibility, and discuss the implications of foreign-accentedness for L2 learners.

2.7 Importance of Intelligibility

2.7.1 Intelligibility of L2 Speech

As Crosby (2020, p. 2) points out, "people want to understand and be understood" in conversation, whether in a serious situation or more informal context. The terms intelligibility and comprehensibility both refer broadly to the ability of a listener to understand L2 speech. Derwing and Munro (1997) and Munro and Derwing (1995), however, define the two terms more narrowly: intelligibility refers to the extent to which an utterance is actually understood, while comprehensibility refers to the native speaker's perception of how easy or difficult it is to understand the message.

Since non-native speech deviates from native-speaker norms on various dimensions, it usually requires more effort on the part of the listener, whether native or non-native (e.g. Brown, McLaughlin, Strand, & Van Engen, 2020; Hendriks, van Meurs, & Usmany, 2021; Hu, Kuo, & Dixon, 2022; Van Engen & Peelle, 2014). For instance, segmental production and the acquisition of phonological contrasts are important for effective communication in an L2, since some segmental errors can hinder listeners' understanding of L2 speech (e.g. Isaacs & Trofimovich, 2012) and can negatively affect L2 word recognition in continuous speech where candidate words are in competition, thereby imposing a higher cognitive load to the listener (Tyler, 2019). While failure to discriminate one or two contrastive pairs may not compromise intelligibility to a large extent, a combination of errors is likely to lead to unintelligible speech (Wells, 2005).

Even though foreign-accented speech does not mean that L2 learners are necessarily unintelligible (Murphy, 2014), several researchers (e.g. Derwing & Munro, 2009; Qian et al., 2018; Suzukida & Saito, 2019; Wells, 2005) point out that the goal of any pronunciation instruction should be to develop sufficiently intelligible speech and increase the mutual intelligibility between speakers, rather than reaching a strictly native-like phonological control. Importantly, pedagogical attention should reflect the fact that not all segments compromise intelligibility to the same extent or in all contexts (Brown, 1988; Qian et al., 2018; Wells, 2005). As Hu et al. (2022) suggest, it is important to identify problematic areas in L2 speech which can impact intelligibility and comprehensibility of speech from learners from specific language backgrounds, in order to allow teachers to prioritise those features in their instruction.

In response to the need to identify and prioritise such problematic areas, some frameworks have been proposed, the most prominent being the Functional Load (FL) principle (Brown, 1988), which provides a list of segmental contrasts ranked based on their communicative value as high FL and low FL categories. According to the FL principle, substitutions of high FL segments lead to lower comprehensibility ratings by native listeners, and therefore, teachers in the language classroom should prioritise teaching them. Munro and Derwing (2006) and Suzukida and Saito (2019) examined how comprehensibility can be impaired by high and low FL segmental substitutions. Both studies demonstrated that high FL errors have more detrimental effects on comprehensibility than low FL errors, while Suzukida and Saito (2019) further showed that high FL consonant substitutions lowered comprehensibility more than high FL vowel substitutions. More specifically, Suzukida and Saito (2019) found that none of the vowels and low FL consonants in their study showed any correlations with comprehensibility, but substitutions of overall segmentals, high FL segmentals (consonants and vowels combined) and high FL consonants were negatively correlated with comprehensibility ratings.

Furthermore, Suzukida and Saito (2019) pointed out that some L2 pronunciation research has shown that L2 vowel errors affect communication less, especially between L2 speakers. While this may initially suggest that vowel errors do not have an effect on comprehensibility when native English listeners are involved, other studies (e.g. Bent et al., 2007) suggest the opposite pattern, i.e. that vowel accuracy, as opposed to consonant accuracy, is correlated with intelligibility. In addition, Thomson (2011) noted that L2 vowels are more difficult to learn, while Neri et al. (2006) argued that the intelligibility of L2 consonants, but not vowels, may improve over time without intervention. This suggests that further research is required to determine the importance of various segmentals on comprehensibility and/or intelligibility for different groups of learners. Given the importance of the L1 on L2 speech production, it seems reasonable to assume that learners' intelligibility and comprehensibility will depend on the production of problematic segments that are determined by the phonemic inventory of their L1. Examining this further has pedagogical implications for learners of specific L1 backgrounds, as it can guide teachers as to which segmentals are important to prioritise in pronunciation instruction and develop an effective pronunciation syllabus.

2.7.2 Implications of Foreign-Accentedness

At the same time, many L2 learners feel that improving their pronunciation is an important goal. For example, 55% of Derwing and Rossiter's (2002) participants responded that at least part of their communication difficulties when speaking English were due to pronunciation problems, while 42% considered these pronunciation problems to be the primary cause of communication difficulties. Similar results were obtained by Derwing (2003), who examined adult immigrants' perceptions of their pronunciation problems and found that most of them not only attributed their problems in communication to their pronunciation, but they felt that a better pronunciation in English would earn them more respect from other people as well.

Accents and dialects are powerful social markers, meaning that they can reveal social information about a speaker to the listener (consciously or unconsciously), who may in turn make inferences about speakers based on their own experiences, attitudes and stereotypical beliefs, sometimes despite their comprehensibility ratings (Baquiran & Nicolaidis, 2020; Carlson & McHenry, 2006; DuBois, 2018; Hendriks et al., 2021). These perceptions may be positive or negative, and may involve factors unrelated to speech, such as social or socioeconomic status, competence, or even intelligence (Cargile, 2000).

Therefore, learners' concerns over pronunciation problems are not unfounded, as pronunciation can entail serious consequences in several aspects of their personal, social and professional life. For example, non-native speakers may face negative reactions based on prejudices (Lippi-Green, 2012) or they may be subjected to accent discrimination including extreme cases of harassment, and refusal of employment or housing (Munro, 2003). In general, foreign-accented speakers tend to receive less favourable ratings than native or standard-accented speakers on both status- and solidarity-related traits (e.g. Dragojevic & Goatley-Soan, 2020).

This has an impact from the perspective of the non-native speakers as well, in relation to anticipated stigmatisation, challenges in communication and feelings of belonging. For instance, Gluszek and Dovidio (2010) examined non-native accented speakers from various backgrounds and found a positive correlation between perceived stigmatisation and difficulties in communicating, as well as a positive correlation between higher levels of perceived stigmatisation and stronger self-rated accents. The researchers also found that having a non-native accent relates to a lower sense of belonging in the United States, both

because of problems in communication and perceived stigmatisation. Furthermore, learners' pronunciation skills in an L2 can negatively affect their self-confidence and social interactions (Gilakjani, 2012). For instance, language proficiency and self-perceived competence in pronunciation in an L2 have been found to be negatively associated with public speaking anxiety (Szyszka, 2011; Tsang, 2022).

Perhaps the most important consequence of foreign-accented speech is the perceived competency it is associated with in relation to jobs and income. Previous research indicates that a person's employability can be affected by their accent (Akomolafe, 2013; Carlson & McHenry, 2006; Derwing & Munro, 2009; George & Chaze, 2014; Roessel, Schoel, Zimmermann, & Stahlberg, 2019) and that foreign-accented speakers may be judged as less competent for high-status jobs, and rated as more suitable for low-status jobs (e.g. Hosoda, Nguyen, & Stone-Romero, 2012). Furthermore, interviewers may favour a group of speakers based on their accent, while employment or advancement might be hindered for speakers with certain accents. This is particularly true for jobs that require strong communication skills (Carlson & McHenry, 2006).

In healthcare, for instance, Baquiran and Nicolaidis (2020) found that foreign-accented doctors may face biases and be perceived as less competent by patients, whether they share the same or different ethnicity. A similar pattern is observed in the domain of higher education, where non-native students tended to evaluate non-native, foreign-accented English lecturers less positively than native English lecturers in various aspects including comprehensibility (Hendriks, van Meurs, & Hogervorst, 2016), attributions of status, likeability, teaching quality (Buckingham, 2014), and competence (Hendriks et al., 2021; Hendriks, van Meurs, & Reimer, 2018).

Importantly, some studies show that having a slight accent is viewed as similar to a native accent (e.g. Carlson & McHenry, 2006; Hendriks et al., 2021). For example, Carlson and McHenry (2006) found that when the speaker's perceived accent was minimal, employability ratings were not affected, while maximally perceived accents yielded lower employability ratings. Furthermore, Hendriks et al. (2021) found that poorer comprehensibility negatively influenced listeners' evaluations of moderately-accented lecturers, but not of slightly-accented lecturers. In addition, some studies found that stronger non-native accented speakers were evaluated less positively than slightly-accented or native speakers (e.g.

Dragojevic, Giles, Beck, & Tatum, 2017; Said, 2006), demonstrating that strength of foreign accent is an important factor in evaluations of non-native speakers, and that there is a general tendency for stronger accents to elicit more negative evaluations (Dragojevic et al., 2017).

At the same time, not all foreign-accented speech receives equally negative evaluations. Such language attitudes are mainly influenced by two underlying processes: social categorisation/stereotyping and processing fluency (Dragojevic & Goatley-Soan, 2020). For instance, in Dragojevic and Goatley-Soan (2020), speakers of some varieties (e.g. Western European) were evaluated more favourably compared to others (e.g. Arabic, Farsi, Vietnamese), based on stereotypes toward different national outgroups. The researchers also found that listeners' attitudes toward foreign-accented speakers also varied as an effect of processing fluency, since not all foreign accents disrupted their processing fluency to the same extent. More specifically, the easier the speech was to understand and process, the more favourable the evaluations were on status and solidarity traits (Dragojevic & Goatley-Soan, 2020). Ease or difficulty in processing fluency may have been due to the phonological features of the accents themselves, the listeners' familiarity with and exposure to the accents, or both.

After reviewing the literature in relation to the concepts of intelligibility and comprehensibility of L2 speech and the possible implications of foreign-accentedness for L2 learners, the following section will discuss the role and beneficial effects of HVPT in L2 segmental acquisition, as a means to move towards a more practical approach to improve L2 instruction.

2.8 High Variability Phonetic Training (HVPT)

Around the 1980s, interest was developed in the role of phonetic variability in speech processing, with researchers shifting their attention to the issue of whether learners could be trained to improve their perception of L2 sounds (Barriuso & Hayes-Harb, 2018). This led to the development of HVPT, a widely researched approach that has been found to improve the perception of both consonants (Carlet, 2019; Hazan et al., 2005; Lively, Logan, & Pisoni, 1993; Logan, Lively, & Pisoni, 1991; Pruitt, Jenkins, & Strange, 2006) and vowels (Carlet, 2019; Iverson & Evans, 2009; Iverson, Pinet, & Evans, 2012; Nishi & Kewley-Port, 2007, 2008; Wang & Munro, 2004). In addition, the improvement achieved was found to be retained several months after the training (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Iverson & Evans, 2009; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994), and it was found to transfer to

speech production for both consonants (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Hazan et al., 2005) and vowels (Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005; Lengeris & Hazan, 2010). Furthermore, the increased variability of input offered in HVPT was found to have a positive effect in the generalisability of knowledge to other speakers and contexts, as it helps learners identify which phonetic cues are relevant to particular sound categories (Thomson, 2011; Wang & Munro, 2004).

The first study investigating whether variability in training input would benefit learners in the perception of L2 contrasts was conducted by Logan et al. (1991), who trained Japanese learners on the English /r/-/l/ contrast and demonstrated that the perceptual categories of adult learners can be altered since learners improved both in trained and untrained stimuli produced by a new speaker. In 1993, the researchers replicated their study with the addition of a generalisation task including both new words and a new speaker (Lively et al., 1993) and demonstrated that new knowledge acquired was generalised beyond training stimuli and training voices, indicating that variability in input leads to robust category formation. Lively et al. (1994) then examined the benefits of HVPT in the long-term retention of knowledge, and found that the improvements were retained at three months, while an insignificant decline was observed at six months; it should be noted, however, that while learners' six-month scores did not differ significantly from their immediate post-test score, they did not differ significantly from their pre-test scores either. Later, Bradlow et al. (1997) further examined the effects of HVPT on production gains and found a clear improvement in participants' productions at post-test. The improvement reported in Bradlow et al. (1997) in both perception and production was later found by Bradlow et al. (1999) to be retained three months later. These studies formed the beginning of the investigation into the effects of this training procedure, which has now "established itself as a major field of inquiry in the speech sciences" (Barriuso & Hayes-Harb, 2018, p. 180).

Since then, various other studies have shown the effectiveness of HVPT in improving segmental perception. For example, Wang and Munro (2004) found that identification training with feedback was effective in improving Mandarin and Cantonese learners' perception of the English vowel contrasts /i/-/ɪ/, /u/-/ʊ/ and /e/-/æ/. In addition, while vowels were initially incorrectly identified based on their quantity, trainees improved in identifying vowels accurately based on their quality instead, similarly to native speakers. Furthermore, Carlet and

Cebrian (2014) examined the effectiveness of a short-term HVPT method on the perception of two English consonant (/v/-/b/ and /d/-/ð/) and two English vowel contrasts (/i/-/ɪ/ and /æ/-/ʌ/) by Spanish/Catalan bilingual learners of English and found a global improvement in the perception of these English vowels and consonants by the trainees.

The following sections provide more details concerning the beneficial effects of HVPT, i.e. the transfer of perceptual gains to production, the generalisation of knowledge to new speakers and contexts, and the retention of the improvement achieved through training.

2.8.1 Production Gains

Perceptual phonetic training and HVPT in particular has been found to help L2 learners create new phonemic categories in their mental representations, and several previous studies (e.g. Bradlow et al., 1997; Huensch & Tremblay, 2015; Jügler, Zimmerer, Möbius, & Draxler, 2015; Lambacher et al., 2005; Lengeris & Hazan, 2010; Motohashi-Siago & Hardison, 2009; Okuno & Hardison, 2016; Rato & Rauber, 2015; Shinohara & Iverson, 2015; Thomson, 2011; Wong, 2013, 2015) have shown that perceptual gains from HVPT can be transferred to production, in line with the perception-production link supported by previous studies (see 2.5.1).

Overall, the examination of studies in Sakai and Moorman's (2018) meta-analysis suggests that perception-only training can yield medium-sized gains for perception and at least small production gains, meaning that it is possible for perception training alone to yield a small but robust improvement in the production modality. At the same time, there is at least a small positive relationship between perception and production gains, providing evidence that the two modalities are connected through the mental representation. According to the researchers, "theoretically, perception training informs the mental representation to become more target-like, which leads to improvements in production" (Sakai & Moorman, 2018, p. 213). If this is the case, then it can be expected that perceptually trained learners should next be able to produce the target sounds more accurately as well (Sakai & Moorman, 2018). This is encouraging, especially for L2 teachers and learners, since time spent training in one modality is beneficial for the improvement of both modalities (Sakai & Moorman, 2018; Thomson, 2011).

However, previous findings in relation to the effectiveness of perceptual training on production gains are inconsistent. While some studies report significant, moderate or partial improvement (e.g. Bradlow et al., 1997; Iverson et al., 2012; Lopez-Soto & Kewley-Port, 2009,

respectively), others found no improvement in production at all, demonstrating a weak or no relationship between the two modalities (e.g. Aliaga-Garcia & Mora, 2009; Garcia Perez, 2003; Peperkamp & Bouchon, 2011), making it difficult to reach robust conclusions as to whether perception training can automatically lead to production improvement.

For instance, Lambacher et al. (2005) trained Japanese learners on five English vowels, and found that after only six 20-minute HVPT sessions, trainees' pronunciation improved, as shown by acoustic analyses and by the intelligibility ratings of 26 native-speaker judges. In addition, Shinohara and Iverson (2015) found that Japanese children's productions of the English /l/-/r/ contrast significantly improved after perceptual training. Thomson (2011) also demonstrated the potential of HVPT to improve speech intelligibility on L2 English vowels without explicit pronunciation training.

Studies including SMG and CYG learners have found similar encouraging results for the beneficial effect of HVPT on the production of English vowels as confirmed through acoustic analyses and perceptual judgments. More specifically, while the SMG speakers in Lengeris (2009a, 2018) and Lengeris and Hazan (2010) used their five L1 categories in their L2 productions before training, the overlap between the English vowels was much less after the training. This indicates that following the perceptual training, these participants learnt to differentiate English vowels in their speech production as well. Lengeris (2009a) and Lengeris and Hazan (2010) noted that L2 vowel perception and production were aligned, but only after training, i.e. after participants were exposed to large amounts of L2 input produced by various native English speakers, providing support for the link between perception and production.

In relation to CYG learners, the only examination of the effect of HVPT on the perception and production of L2 English vowels was conducted by Georgiou (2021), who trained CYG children and adults with little experience in English on the full set of RP vowels. Georgiou (2021) administered five training sessions consisting of 220 stimuli each, over a period of three weeks, and found that HVPT was effective in improving trainees' identification accuracy, particularly for children who showed greater gains. However, transfer of perceptual gains to production was only significant in children and not adults, suggesting that children produced the vowels more accurately than adults at post-test, according to the identification scores of three native English speakers.

On the other hand, Aliaga-Garcia and Mora (2009) investigated whether a six-week phonetic training paradigm could improve accuracy in the perception and production of the sound contrasts /p/-/b/, /t/-/d/, /i:/-/ɪ/ and /æ/-/ʌ/ of advanced Catalan/Spanish bilingual learners of English. The researchers found significant differences in the discrimination scores of the experimental group before and after training for the vowel contrasts, but no significant improvements were observed in L2 vowel production as a result of phonetic training. However, even though no overall significant gains in perceptual or productive competence were observed in this study for all the sound pairs examined, there was significant improvement in either perceiving or producing some of the target sounds examined after the training. Overall, this study demonstrates that phonetic training can have different effect sizes on learners' perceptual and productive competence based on phonetic dimension and sound contrast. Furthermore, Zhang, Cheng, Qin, and Zhang (2021) found no significant improvement on the intelligibility of adult Mandarin Chinese learners of English on the production of the English /i/-/ɪ/ contrast after HVPT, even though the researchers noted a significant gain in the use of spectral cues and decrease in the use of the secondary durational cue.

These studies suggest that even though accurate production may not always necessarily depend on accurate perception, perceptual training may lead to improvements in segmental production, without any production training. The following section will focus on the effects of HVPT on the generalisation of knowledge to new speakers and contexts.

2.8.2 Generalisation

In addition to its benefits for production gains, previous studies suggest that the increased variability within a phonetic category offered in HVPT can promote the generalisation of perceptual gains from familiar to new voices and from trained to new words, thereby demonstrating that any positive effect of the training can be transferred beyond the training stimuli (e.g. Brosseau-Lapré, Rvachew, Clayards, & Dickson, 2013; Carlet & Cebrian, 2014; Carlet, 2019; Cebrian et al., 2019; Iverson, Hazan, & Bannister, 2005; Kondaurova & Francis, 2010; Lively et al., 1993; Qian et al., 2018; Sadakata & McQueen, 2013; Thomson, 2011; Wang & Munro, 2004; Wong, 2012, 2014). This is arguably due to exposure to highly variable input (but see 2.9.1), which encourages learners to form more generalised representations of a sound, as it helps them identify which phonetic cues are relevant to particular sound

categories and exclude any irrelevant, speaker-identity cues, thereby developing a more native-like cue weighting (Aliaga-Garcia & Mora, 2009; Carlet & Cebrian, 2014; Giannakopoulou, Brown, Clayards, & Wonnacott, 2017; Thomson, 2011) (see 2.5.4).

This transfer of knowledge from one setting to another is of particular importance, as "successful transfer is integral to robust learning" (Qian et al., 2018, p. 76), which is perhaps the most important outcome of HVPT (Logan & Pruitt, 1995). Since speech variability tends to interfere with learners' perception capacity, assessing transfer from familiar to new voices is important given that the training can be considered successful only when trainees are able to accurately perceive trained sounds produced by unfamiliar voices (Qian et al., 2018). Assessing transfer to new words is also important because the perception of speech sounds is context-dependent, i.e. the acoustic characteristics of a phoneme can be affected by its surrounding phonetic environment (allophonic variation) (Flege, 1995; Qian et al., 2018; Strange, 2007; Thomson & Derwing, 2016; Thomson & Isaacs, 2009; Thomson, 2012); therefore, learning in one phonetic context does not mean automatic improvement in others and additional experience with various phonetic environments is necessary for their formation and generalisation of knowledge to new contexts (Logan et al., 1991).

Previous studies assessing generalisation of learning to new contexts and speakers present encouraging results, although they vary in terms of the extent of generalisation. Some studies show that learners can generalise their newly acquired knowledge to new contexts and/or speakers (e.g. Lively et al., 1993), while others show only partial generalisation (e.g. Iverson et al., 2005). However, a general tendency for positive results has been observed, especially in comparison with low variability phonetic training (LVPT). For example, Giannakopoulou et al. (2017) compared high- and low-speaker variability training in L2 segmental learning with native SMG adults and children and although they did not find robust evidence that HVPT is more beneficial than LVPT, they report a tendency for greater improvement in the HVPT condition, especially for adults. The lack of a statistically significant difference between the two conditions may have been the result of the overall high performance of adults, who hit ceiling values in the LVPT condition.

Furthermore, the Mandarin learners of Canadian English in Thomson (2011) showed improvement in their intelligibility of Canadian English vowels that were produced in an elicited imitation task that included both a known and a new voice, as well as improvement in

the production of vowels in some new phonetic environments (/zV/ and /sV/), but not others (/gV/ and /kV/). This was attributed to the similarity in the articulatory position of the trained /b-p/ contexts and the /z-s/ contexts, as opposed to the /g-k/ contexts, as transitions from this pair into the following vowel would be different and would therefore lead to differences in perception. These results indicate that training helped learners isolate the relevant phonetic cues to vowel identity, and that these were then generalisable to new speakers.

More recently, Qian et al. (2018) found that in addition to significant improvements in the identification and discrimination of target phonemic contrasts on trained items, the training was also effective in improving participants' ability to generalise their perceptual discrimination and identification abilities to new, untrained voices. However, the participants in this study failed to discriminate and identify trained phonemic distinctions in new words, suggesting that the training was not successful in facilitating the transfer of perceptual gains from trained to untrained words. The researchers suggested that intensifying the training through longer sessions and more enriched training stimuli with a variety of phonemic variations could have enhanced generalisation to new contexts.

2.8.3 Retention

Another important aspect of HVPT is the retention of learning, which has been found to last for several months (Carlet, 2017; Cebrian et al., 2019; Iverson & Evans, 2009; Lively et al., 1994; Rato & Rauber, 2015; Thomson, 2012; Wang & Munro, 2004; Wang, Spence, Jongman, & Sereno, 1999), suggesting that HVPT can help learners make long-term changes in the way that new sound categories are represented in memory. Previous studies have assessed the retention of learning at various stages after the training, with promising results. For instance, Thomson (2012) and Rato and Rauber (2015) included delayed post-tests and found that the improvement achieved by the experimental group was retained one and two months after the training, respectively. Carlet (2019) also found that Spanish/Catalan trainees' perception of English vowels remained similar between the immediate and the delayed post-test two months later, and both were significantly improved compared to the pre-test, while the performance of the control group remained similar across all three times, an indication that robust learning had taken place for the experimental group.

Longer periods of retention were observed in Wang and Munro (2004), Cebrian et al. (2019) and Wang et al. (1999). Wang and Munro (2004) found no significant decline in

participants' performance in any of the target vowel pairs in a retention test three months after the training; even though trainees' performance was slightly lower than in the post-test, it was still substantially better than in the pre-test. Cebrian et al. (2019) found that the improvement was retained four months later, as indicated by the lack of a significant difference between post-test and delayed post-test results. Finally, Wang et al. (1999) provided evidence that perceptual learning can be retained without a decline for at least six months; however, their study included training of Mandarin tone contrasts, and therefore, it is yet unknown whether retention for this period of time is possible in the training of segmental categories (Thomson, 2018).

Assessing retention is important, as it shows that learners have not only acquired the muscular control required for the production of L2 segments, but also learnt to turn this muscular control into a habit. According to Koutsoudas and Koutsoudas (1962), the former is easier, as conscious effort may lead learners to the use of the correct movements. Learning to perceive and produce sounds with conscious effort without acquiring the new habit, however, may not lead to automatic production, since in this case learners are likely to continue to use familiar habits in the production of L2 sounds (Koutsoudas & Koutsoudas, 1962). However, despite its importance, the use of delayed post-tests to assess the retention of learning is not very common in previous studies (Sakai & Moorman, 2018; Thomson, 2018).

Having reviewed the benefits that an HVPT paradigm can offer to L2 learners, the following section will focus on some of the methodological choices employed in HVPT studies, and how these may affect the success of the training.

2.9 Methodological Considerations in HVPT Paradigms

It is of crucial importance to note that HVPT studies vary widely in terms of methodological choices and training paradigms used, which can play a key role on the outcomes on perceptual and production performance and can explain, at least partly, the inconsistency in the results found in previous studies, as pointed out by many researchers (e.g. Barriuso & Hayes-Harb, 2018; Hu et al., 2022; Melnik-Leroy et al., 2022; Nagle & Baese-Berk, 2022; Sakai & Moorman, 2018; Thomson, 2018). As Nagle and Baese-Berk (2022) argue, the diverse findings in the literature reflect the diverse methodological choices that researchers have made, including, among others, the length of the training, the task type and complexity and the analysis of the results. Furthermore, differences in the results may also be attributed to individual differences

among participants, L2 experience, L1 background, training methods used and the nature of the acoustic variability in the input (Barriuso & Hayes-Harb, 2018; Kartushina & Martin, 2019). Thus, the extensive research conducted on HVPT and its effects also includes investigations that aim to identify the methodological approaches that are best for optimal learning (see Lee, Jang, & Plonsky, 2015 for a review).

The following sections aim to highlight two important aspects of HVPT manipulated in previous studies in an attempt to establish effective practices, i.e. stimulus variability and the use of modified input.

2.9.1 High vs. Low Variability in Stimuli

The beneficial effects of HVPT, particularly in terms of the generalisation of knowledge, have long been largely attributed to the exposure it offers to variable input, including multiple talkers and multiple phonetic contexts, which are generally considered as a requirement in order for the training to be effective. Furthermore, in addition to generalisation of learning to new words and speakers, the superiority of HVPT over LVPT has also been reported in relation to an improved perceptual performance and transfer of perceptual learning to production (Brosseau-Lapré et al., 2013; Kartushina & Martin, 2019; Sadakata & McQueen, 2013; Wong, 2012, 2014; Zhang et al., 2021).

According to Thomson and Derwing (2016), since allophonic variation is natural in speech, pronunciation instruction should incorporate such variation as well. In addition, Thomson (2011) argues that training with a single voice may direct learners' attention towards any cues (generalisable or not) that can help them discriminate between training stimuli, instead of focusing on relevant cues. At the same time, learners have been shown to demonstrate speaker variability in their ability to perceive speech, meaning that they might be able to perceive speech from one voice better than from another, and are usually better able to perceive speech from a voice they have been trained on rather than an unfamiliar voice (Wang & Munro, 2004). The highly variable stimuli offered in HVPT, both in terms of speaker and context, can therefore help learners identify and focus on those acoustic cues that are relevant for accurate identification and discrimination, and ignore speaker-specific, irrelevant cues (Brekelmans, Lavan, Saito, Clayards, & Wonnacott, 2022; Kondaurova & Francis, 2010).

However, although earlier research strongly supported the use of multiple speakers as a source of variability in HVPT, some studies have yielded mixed results as to the benefits of talker variability. For instance, even though greater talker variability was found to facilitate generalisation further, single-talker training was also found to be effective in the generalisation of knowledge (Perrachione, Lee, Ha, & Wong, 2011; Wong, 2012, 2014), while some studies support that single-talker training yields the same effects in terms of generalisation (Dong, Clayards, Brown, & Wonnacott, 2019; Giannakopoulou et al., 2017). Furthermore, Brekelmans et al. (2022) replicated the studies conducted by Logan et al. (1991) and Lively et al. (1993) with a larger sample size, to examine whether HVPT is superior to LVPT. The researchers found that trainees performed better after training irrespective of the type of training (high- or low-variability) and concluded that if HVPT is more beneficial, this effect might not be as large as previously assumed.

The fact that some studies found that single-talker training can also lead to generalisation of learning suggests that there may be other sources of variability in the single-talker condition that may also facilitate generalisation (Zhang et al., 2021). To this end, Zhang et al. (2021) examined whether talker variability is indispensable for generalisation, by comparing multiple-talker training with two types of single-talker training, one with irrelevant acoustic variability and audio-visual input and one without these “enhancement” features. The researchers found that both the multiple-talker training and the single-talker training with the additional features yielded similar results, helping Mandarin Chinese learners improve in the identification of the target English vowels /i/-/ɪ/ and generalise their knowledge to new speakers and contexts, although both groups failed to improve their intelligibility; however, when the additional features of audio-visual input and adaptive acoustic exaggeration were removed, there were significant advantages of the multiple-talker over the single-talker training paradigm. This study shows that talker variability may be unnecessary when enhanced acoustic variability along a secondary dimension is included, since the latter can also induce robust learning.

In addition, some studies have observed that variability can also hinder learning, at least in some cases. For example, variability was found to impede learning when the target L2 contrasts were difficult or highly confusable in relation to the learners’ L1 (e.g. Giannakopoulou et al., 2017; Wade, Jongman, & Sereno, 2007), when the target group

involved children or young learners (e.g. Evans & Martín-Alvarez, 2016; Giannakopoulou et al., 2017), or when the target group involved perceptually weak or novice learners, due to the added processing costs required to process speech by multiple rather than a single talker (Antoniou & Wong, 2015; Chang & Bowles, 2015; Perrachione et al., 2011; Sadakata & McQueen, 2014).

Furthermore, some studies suggest that variability in perceptual training may have different effects when it comes to production. For instance, Kartushina and Martin (2019) found that even though training was effective in improving novice learners' production of both target vowels (French /e/-/ɛ/) in both conditions (single-talker and multiple-talker), production improvement was higher in the single-talker than in the multiple-talker condition. The researchers attributed this to the increased cognitive load required to learn two challenging vowels while managing the variability from multiple talkers, as opposed to learning them with a single, familiar voice, particularly in the case of inexperienced learners. In addition, even though Brosseau-Lapré et al. (2013) reported that only multiple-talker training promoted generalisation to an unfamiliar speaker, they found no significant differences in the amount of production improvement in their participants between the multiple-talker and single-talker conditions or the high- and low-variability stimuli conditions. Evans and Martín-Alvarez (2016) on the other hand, found that LVPT was more effective than HVPT in promoting production improvements.

2.9.2 Modified Input

In addition to variability in input, researchers have also examined whether the use of modified input for cue enhancement or cue inhibition is more beneficial than using naturalistic stimuli. Adaptive training for cue enhancement involves starting with clearly distinguishable stimuli with exaggerated values, and gradually moving to stimuli with reduced perceptual difference, so that perceptual acuity improves over the course of training (Kondaurova & Francis, 2010). Inhibitory training introduces irrelevant variability along the more-attended dimension, in order to encourage listeners to ignore it in categorisation (Kondaurova & Francis, 2010). The use of modified stimuli in such training paradigms removes the reliability of the durational cue in the input, thereby forcing learners to attend to spectral cues to successfully discriminate the vowels in the contrast (Kondaurova & Francis, 2010; Yuan & Archibald, 2022).

Kondaurova and Francis (2010) examined the effects of cue enhancement and cue inhibition on the acquisition of the American English tense-lax contrast /i/-/ɪ/ by Spanish learners and found that while all learners relied on vowel duration initially, they all increased their reliance on spectrum properties at post-test, irrespective of the training type. The researchers noted, however, that inhibitory training was more effective compared to enhancement training, while both types of training were more effective compared to training with a natural cue distribution. Similar results are reported by Ylinen et al. (2010), who found that Finnish learners of English improved in their ability to focus on spectral differences in identifying the vowels in the /i/-/ɪ/ contrast after identification training with both natural and synthesised stimuli. Hu et al. (2016) used training with duration-equalised vowels, thus making duration an unreliable cue to vowel perception, and also found that it helped Mandarin Chinese learners reduce their reliance on duration in the perception of L2 vowels.

More recently, Cheng, Zhang, Fan, and Zang (2019) and Yuan and Archibald (2022) found equally encouraging results. More specifically, Cheng et al. (2019) compared the effects of HVPT with temporal acoustic exaggeration to a typical HVPT paradigm on Chinese learners' perception of the English /i/-/ɪ/ contrast and reported that both training types significantly improved learners' identification of naturally produced words by new speakers, while the group trained under the exaggerated condition showed greater improvement in natural word identification in words produced by new speakers and target vowels produced in new contexts. Finally, Yuan and Archibald (2022) examined Mandarin Chinese EFL learners and whether HVPT with modified input can help them re-weight their perceptual cues to attend to spectral differences in the discrimination of English /i/ and /ɪ/. The researchers found that the experimental group improved significantly more than the control group that had exposure to naturalistic input alone.

These studies provide evidence that using modified input can enhance EFL learners' perception of challenging L2 contrasts and facilitate phonetic learning, suggesting that it should be incorporated into the HVPT paradigm.

2.10 Summary

This chapter has focused on a review of the literature in relation to the linguistic situation and the status of English in Cyprus, and presented the differences between the simple 5-vowel system of SMG and CYG compared to the more complex vowel system of English, and

specifically SSBE. As explained within the framework of current theoretical models such as the SLM, the SLM-r, the PAM and the PAM-L2, these differences may explain the difficulties that learners encounter in perceiving and producing the members of an L2 phonological contrast. Furthermore, the chapter also discussed other possible factors that may have a role to play in L2 segmental acquisition, focusing on language use patterns, motivation and input. Possible implications of foreign-accentedness such as negative evaluations, stereotyping and its effects on perceived competency have also been discussed, highlighting the importance of understanding those factors so that the appropriate tools are developed for effective L2 pronunciation instruction. Such tools can include the HVPT paradigm, as discussed in sections 2.8 and 2.9, a popular and seemingly effective technique of pronunciation teaching, which, however, necessitates further research in order to optimise its implementation and results.

The following chapter deals with the methodological considerations and approaches taken to address each RQ, including a detailed description of the participants, procedures, tools and tasks. The chapter provides information about the HVPT paradigm and testing procedures used in the study, and finally, it describes the quantitative and qualitative data analysis procedures followed.

CHAPTER 3: METHODOLOGY

3.1. Research Questions and Methods

As shown in the previous chapter, the acquisition of L2 segments can pose considerable difficulties for learners, and depends on various factors, both extrinsic and intrinsic to the learners. Current models of speech perception and production as well as previous research in the field suggest that the differences between the two vowel systems under investigation are likely to add to the difficulties that CYG learners of English face in L2 vowel perception and production. At the same time, individual differences between participants can also affect acquisition. As discussed in 2.7, foreign-accentedness may have a significant impact on learners' personal and professional lives, and therefore, research investigating the possible factors involved in L2 segmental acquisition are crucial in order to better understand and use them in L2 instruction. Furthermore, as demonstrated in 2.8, the HVPT paradigm can be a promising tool for L2 instruction, although its implementation and effects vary across studies, yielding contradictory results. The present study aimed to further investigate these factors by examining the L2 perception and production of English vowels by an understudied population through the following RQs:

1. How do CYG adult learners perceive and produce L2 English vowels?
2. What are the spectral and durational differences in English vowel production between CYG learners and NE speakers?
3. How effective is HVPT in improving CYG learners' vowel perception and production without explicit production training?
4. How well does any improvement generalise to new speakers and contexts, and to what extent is it retained after a two-month period?
5. Do individual differences in motivation, input and language use patterns affect CYG learners' perception and production of L2 English vowels?

A quasi-experimental design was followed for RQs 1-4, as participants were purposefully selected to fulfil certain predefined criteria and the tasks were completed under controlled circumstances (Mildner, 2013). More specifically, this was a non-equivalent groups design involving participants from two populations (NE speakers and CYG learners) carefully selected so as to share as many characteristics as possible (both within and across groups), especially

in terms of age and educational background (Mildner, 2013). Furthermore, a non-equivalent control-group design was followed, as CYG participants were not randomly selected or assigned to the experimental and the control groups, the treatment was provided to the experimental group only, and both groups completed a pre- and a post-test to ensure that the outcome was a result of the intervention and not any other factor (Creswell, 2014). Although unknown or inevitable differences between groups can pose threats to the internal validity of the study, such as selection bias, this design can control for other threats to internal validity, such as instrumentation, which remains constant across groups (Creswell, 2014; Mildner, 2013).

As concerns RQ 5, both quantitative and qualitative data were collected in order to reach a comprehensive understanding of the issue under examination. An explanatory sequential mixed methods design (QUAN→qual) was followed, as the quantitative data were used to purposefully select individuals for the qualitative phase and identify the open-ended questions of the interviews (Creswell, 2014). Quantitative data were obtained following a non-experimental correlational design in order to examine whether there was a relationship between the three independent variables (IVs), i.e. language use patterns, motivation and input, measured through a questionnaire, and the production and perception scores of CYG learners as obtained for RQs 1-4 (Mligo, 2016). However, given the limitations of questionnaires in measuring these variables (Flege, 2008), qualitative data in the form of interviews were collected as well, using criterion sampling. More specifically, extreme cases were interviewed so that complementary information could be obtained (Sandelowski, 2000).

3.2 Participants

Fourteen CYG (7 male, 7 female) and 10 NE speakers (4 male, 6 female) were recruited for the purposes of this study. Four additional CYG participants chose to withdraw from the study after completing the pre-test, and therefore, all the data collected from them were deleted. All participants were volunteers recruited via email invitations, advertisements, distribution of flyers and the word of mouth, and the procedure adhered to all ethical standards. None of the participants reported having a speech or hearing impairment. In order to ensure that the groups were as homogeneous as possible and the results comparable to other studies (Palinkas et al., 2015), participants needed to fulfil a set of criteria. The criteria for CYG participants involved being native speakers of the CYG dialect between the ages of 18 to 28

years old, born and raised in a CYG-speaking community and studying at an English-speaking university in Cyprus. Similar criteria were set for NE participants, who had to be native speakers of Standard Southern British English (SSBE), between 18 and 28 years old, and study at an English-speaking university in the UK. The target variety chosen was SSBE, as this is an extensively examined variety preferred in previous studies (e.g. Aliaga-Garcia, 2017; Escudero & Chladkova, 2010; Giannakopoulou, Uther, & Ylinen, 2013a; Krzonowski et al., 2015; Lengeris 2009a; Lipińska, 2017, among others), and it is considered to be a form of reference speech, commonly used as the pronunciation model in various parts of the world (Deterding, 1997).

The CYG participants in this study were 18-25 years old (average: 20.5) and they were students of various disciplines in various years of study (see **Table A1** in Appendix A for further information). A restriction was placed for English Language and Linguistics students, who had to be in their 1st year of studying, before attending any phonetics or phonology modules; only one student who fulfilled this criterion was recruited. Participants had not studied or lived abroad for extensive periods of time, i.e. more than 1 month. All reported that CYG was their native language, their parents' native language and that they had been born and brought up in a CYG community where the main language used was CYG. They had started learning English at an average age of 9 years old (range: 7-12 years old) and had been learning English for an average of 8.6 years (range: 6-11 years). They had all attended Greek-speaking schools in Cyprus during their primary and secondary education and they reported having learnt English mostly through afternoon lessons (14 responses) and through movies or series (8 responses). Based on self-reports of participants' English language skills, CYG learners rated their understanding at 4.9/6 (SD = 0.8), their speaking skills at 4.1/6 (SD = 1.2), their writing skills at 4.4/6 (SD = 1) and their listening skills at 4.4/6 (SD = 1.3).

CYG participants reported that they used their native language all the time at home and most of the time in social settings during their school years, although some reported using English sometimes in social settings to communicate with non-native speakers (e.g. at work, when abroad, or in other social contexts). During their university years, they all reported using English in class all the time or almost all the time, while mostly using CYG at home and in social settings. Using (speaking or listening to) English at home and in social settings was also reported by participants to various extents. Based on the information provided, they were determined to be intermediate to advanced EFL learners of English, who used their L2 on a

regular basis. None of the CYG participants had attended any phonetics or phonology modules or had completed any pronunciation training in the past.

NE speakers were 19- to 28-year-old students at an English-speaking university in the UK studying in any discipline (see **Table A2** in Appendix A for further information), forming a group comparable to the CYG group and limiting the effects of age or educational background as much as possible. In addition to these, NE participants were born and raised in England, used English as their main language at home and in school, and came from a single regional background, thus limiting regional variation effects to the minimum (Palinkas et al., 2015). All NE speakers reported that they believed they had an SSBE accent, most were brought up in a southern region of England, and most reported that they did not speak another language above an intermediate level. However, some participants reported living in various other countries or regions of England for more than 12 months, speaking other languages to more advanced levels, or having a parent speaking a different language or variety of English. This variation is expected when collecting naturalistic data and has been accepted in previous studies as well (e.g. Kondaurova & Francis, 2010); finding "ideal" participants would be an unrealistic expectation, especially considering the multilingual and multicultural nature of our communities nowadays (Douglas Fir Group, 2016). Nevertheless, due to the sensitive nature of phonetic data, this variation was limited in the study through the rating of NE participants' recordings by 2 NE speakers trained in phonetics and 3 NE naïve listeners, as described in 3.5.

3.3 Procedure

Before completing any tasks, all participants filled in a linguistic background questionnaire (see Appendix B.1 and B.2 for the questionnaires completed by CYG and NE participants, respectively), providing information about their linguistic and educational background. In addition to confirming that participants fulfilled the criteria of inclusion in the study, this questionnaire also served as the main tool used to address RQ 5, i.e. the effect of language use patterns, motivation and input on the performance of learners. More specifically, the questionnaire included items for the assessment of participants' L1-L2 use patterns during secondary and higher education, motivation levels, overall exposure to the L2 and quality and quantity of L2 input received during secondary and higher education. In order to ensure its clarity, the questionnaire was pilot-tested with 10 individuals matching the potential participants in age and educational background. The questionnaire was sent to the

participants via email to be completed before their first session (pre-test), although they were encouraged to request clarifications at any point before or during the first session. The questionnaire was translated in Greek and participants had the option to choose their preferred version.

In addition to the linguistic background questionnaire, NE speakers only completed the production task described in 3.6. The procedure followed by CYG participants is outlined in **Table 1**. More specifically, CYG participants completed a pre-test, an immediate post-test and generalisation test three weeks later, and a retention test two months later. All three tests included both perception and production tasks; the perceptual task (described in 3.5) was performed before the production tasks (described in 3.6) in order to avoid recognition effects, as in Detey and Racine (2015). In addition, as in Lengeris and Hazan (2010), all tasks with English vowels preceded those with Greek vowels. The stimuli used are described in 3.4. Eight out of the 14 CYG participants also completed a set of 8 training sessions (experimental group), while the remaining 6 participants formed the control group, to examine the effects of the training. Furthermore, trainees also completed a post-training questionnaire (see Appendix B.3) evaluating the training provided, and three of them were invited to qualitative interviews.

PHASE	TASKS
PRE-TEST (T1)	1. Linguistic Background Questionnaire 2. Perceptual Test: Forced-Choice Identification Task (FCID) without feedback 3. Production Task: Elicited imitation task 4. Production Task: English wordlist-reading 5. Production Task: Cypriot-Greek wordlist-reading
TRAINING (8 SESSIONS)	HVPT: FCID with feedback
POST-TEST (T2)	1. Post-training evaluation questionnaire (trainees only) 2. Perceptual Test: FCID without feedback (identical to pre-test) 3. Generalisation test 4. Production Task: Elicited imitation task 5. Production Task: English wordlist-reading
RETENTION TEST (T3)	1. Perceptual Test: FCID without feedback (identical to pre-test) 2. Generalisation test 3. Production Task: Elicited imitation task 4. Production Task: English wordlist-reading

Table 1. Data collection procedure (CYG participants)

Since the tests consisted of a large number of stimuli and they were repetitive in nature, participants were encouraged to take breaks between the perceptual and production tasks, as well as halfway through the stimuli in each task, which were taken as needed. Oral instructions were given in CYG, so that participants would feel more comfortable, in an attempt to elicit naturalistic speech as much as possible (Bernhardt, Bacsfalvi, Adler-Bock, Modha, & Purves, 2013; Hoffman, 2014).

3.4 Stimuli

3.4.1 Natural Stimuli

The stimuli used in the study included the 11 SSBE target vowels in a CVC context and are shown in **Table 2** below. The /bVt/ and /gVt/ stimuli were used in the pre-test, post-test and retention test, as well as the training sessions, whereas the /sVt/ and /dVt/ contexts were reserved for the generalisation test. The CYG stimuli were only produced by the CYG speakers.

	ENGLISH STIMULI					CYG STIMULI	
CONTEXT / VOWEL	bVt	gVt	sVt	dVt	CONTEXT / VOWEL	'bVtV	'pVtV
ɪ	bit	git	sit	-	i	*/'bita/	/'pita/
i:	beat	*gheat	seat	-	e	*/'beta/	/'peta/
e	bet	get	set	debt	a	*/'bata/	/'pata/
ɜ:	*burt	*gert	cert	dirt	o	/'bota/	*/'pota/
ɑ:	*bart	*gart	-	dart	u	*/'buta/	*/'puta/
æ	bat	gat	sat	-			
ʌ	but	gut	subtle	Dutch			
ɒ	bot	got	sot	dot			
ɔ:	bought	*gort	sought	-			
ʊ	butch	good	soot	-			
u:	boot	*gould	suit	-			

Table 2. English (adapted from Bohn & Bundgaard-Nielsen, 2008; Lengeris, 2009a; Mayr & Escudero, 2010) and CYG stimuli (adapted from Papachristou, 2011)

As in other studies (e.g. Bohn & Bundgaard-Nielsen, 2008; Huensch & Tremblay, 2015; Mayr & Escudero, 2010) both real and non-words (indicated by asterisks in **Table 2**) were included in the stimuli in order to ensure that the vowels appeared in matching environments in both languages (Di Paolo, Yaeger-Dror, & Wassink, 2011; Kerswill & Watson, 2014), and to minimise potential word-frequency effects (Huensch & Tremblay, 2015), although real words were preferred when available. Importantly, this also helped in directing the attention of

learners only on phonetic form and not meaning (Thomson, 2011). Thomson and Derwing (2016) assessed the use of real and non-words in perceptual training and found that training with predominantly non-words was superior to training with real words in achieving pronunciation improvement, since it forced learners to attend to relevant phonetic details. The researchers, therefore, supported that L2 vowel instruction should focus on phonetic level information rather than placing exclusive attention to pronunciation in real words. This is particularly important, since adult learners tend to focus on meaning rather than form in natural speech (Schmidt, 2001).

The choice of contexts for the English stimuli was carefully considered and based on previous studies (e.g. Bohn & Bundgaard-Nielsen, 2008; Lengeris, 2009a; Mayr & Escudero, 2010; Papachristou, 2011). The rationale was to maintain contexts sharing the same manner of articulation for the stimuli that would form the training sessions (/bVt/ and /gVt/), and assess the generalisability of learning to real words only, in a context with a consonant of the same manner of articulation (/dVt/) and a context with a consonant of a different manner of articulation (/sVt/) (adapted from Thomson, 2011). The /sVt/ contexts was preferred in the latter case due to the larger number of real words in this context.

As concerns the CYG stimuli, a disyllabic ˈCVCV structure was preferred in order to ensure that they appeared in phonotactically permitted sequences, thereby limiting the difficulty that prohibited structures may entail for learners (Thomson, 2011). Only the stressed vowels were analysed in the CYG stimuli. Furthermore, in order to avoid overwhelming the participants given that CYG vowel analysis was not part of the key purposes of the study, only two contexts were included for the CYG stimuli, corresponding to the English /bVt/ context. Both a /'pVtV/ and a /'bVtV/ context were used so as to address aspiration differences between the two languages. More specifically, SMG and CYG voiceless plosives /p, t, k/ are unaspirated (Arvaniti, 2001, 2007, 2010; Botinis, Fourakis, & Prinou, 2000) and English voiced stops /b, d, g/ are also unaspirated and phonetically realised as voiceless in initial position (Docherty, 1992). Therefore, the CYG /p/ was considered to be a more appropriate counterpart of English /b/ in initial position, as the two sounds are phonetically realised in a similar way. However, since CYG learners were likely to produce the English stimuli in different ways (either with a British accent, which would be closer to the voiceless initial stops in CYG, or with fully voiced stops, i.e. closer to the Greek voiced stops) both the /'pVtV/ and the

/'bVtV/ contexts were included, in order to enable the analysis of vowels in the most appropriate context. During analysis, it was established that in most cases, CYG learners produced the English stimuli with a fully voiced stop, and therefore, comparisons with CYG vowels were made using learners' productions in the /'bVtV/ context. Furthermore, since the consonants included in the CYG contexts (word-initial and intervocalic plosives) could be pronounced as geminates by CYG speakers, these participants were informed that all consonants in these words should be pronounced as singletons.

3.4.2 Synthetic Stimuli

Since it was previously observed that L2 learners of English tend to inappropriately rely on vowel quantity rather than quality to distinguish the members of contrastive vowels (e.g. Cebrian, 2006; Mora & Fullana, 2007; Thomson, 2011; Wang & Munro, 2004), synthetic stimuli were also used in the training sessions, in order to draw trainees' attention away from vowel length and towards vowel quality by making duration an unreliable cue to vowel identification. As noted in 2.9.2, enhancing relevant acoustic cues by modifying the stimuli is associated with increased learning, since the relevant cues are made more salient and learners are more likely to notice important differences between L2 categories (Cheng et al., 2019; Escudero, Benders, & Wanrooij, 2011; Thomson, 2012; Ylinen et al., 2010; Yuan & Archibald, 2022).

Therefore, a Praat (Boersma & Weenink, 2013) v 6.1.37 script was used to create a 5-step duration continuum for each target item (60ms, 120ms, 180ms, 240ms and 300ms) using the original recordings of NE speakers (see Appendix C.1 for the script used). The shortest and longest durations of the continuum were determined through an examination of the recordings of NE speakers and represent the shortest and longest vowel productions of NE speakers across all vowels. Only the vowel portion of each target word was manipulated, without changing any other acoustic information. All synthetic stimuli were checked for naturalness, and distorted items were excluded.

3.5 Perceptual Tasks

Perceptual tasks were developed using the stimuli produced by NE speakers. More specifically, the target words produced by NE speakers were isolated from the carrier phrase "He said…and left" (Recording 1; see 3.6) and imported into the TP (Teste/Treino de Percepção – Perception Testing/Training) v 3.1 software (Rauber, Rato, Kluge, & Santos, 2012) for the preparation of

the training sessions and the perceptual tests (pre-test, post-test, generalisation test) completed by CYG learners.

In an effort to minimise variation as a result of the linguistic background of NE participants and ensure that the vowels were produced in the target variety (SSBE), two NE raters trained in phonetics assessed their quality. Raters were instructed to exclude stimuli that were unclear due to recording or voice quality, stimuli in which the target vowel did not correspond to the target SSBE vowel, or stimuli that revealed any characteristics indicating the use of a different regional accent. A total of 35 stimuli were excluded from the original 370 assessed (10 NE speakers x 37 stimuli). Further to the assessment of the quality of the target vowels by NE raters trained in phonetics, the perceptual tasks were also completed by 3 NE naïve listeners, to ensure their validity and reliability. Listeners achieved an average accuracy rate of 87.86% (Rater 1: 85.07%, Rater 2: 89.55%, Rater 3: 88.96%). Stimuli that were wrongly identified by two or more raters were removed through this process, resulting in 28 items being further removed from the perceptual tasks and all other analyses. Similarly to Thomson (2011), the remaining stimuli that were perceived as an ambiguous exemplar of the target category (misidentified by one listener only) were not considered problematic, as they represent a very small proportion of the stimuli, they reflect the variation found in naturalistic speech, and their misidentification may also have been due to individual listener characteristics.

All perceptual tasks were administered using the TP software, an open-source application software developed for speech perception experiments and perceptual training tasks by Rauber et al. (2012). They all were forced-choice identification (FCID) tasks, in which participants heard each stimulus in isolation through headphones up to 9 times, and then clicked on the label containing the word they heard on a computer screen, choosing from a set of options provided, depending on the task. Each target word with its corresponding label was introduced before each test, but participants were encouraged to read the labels again once the tests started, to ensure that they were familiar with their options. In addition, participants were instructed to ignore differences in speakers' voice and focus on the vowels in each target word (Lengeris, 2009a). An "Oops" button was enabled in all perceptual tests, allowing participants to choose a different label if they accidentally clicked on the wrong button in their immediately previous response.

It should be noted that all occurrences of the target words "butch", "subtle" and "Dutch" were manipulated in all perceptual tasks, so that context would not facilitate the identification of the target word. In these cases, participants only heard /bʊt/, /sʌt/ and /dʌt/, respectively. Their respective labels in the tasks were also modified as follows: the label "bUt" was used for /bʊt/ to differentiate it from "but" (see **Figure 1** in 3.7), the label "sut" was used for /sʌt/ and the label "dut" was used for /dʌt/ (see **Figure 4** in 3.9). In order to avoid confusion, participants were instructed in advance about these labels; during the presentation of labels with their corresponding target words, it was stressed that the label "but" corresponded to the real word as they know it, the label "bUt" corresponded to the vowel found in the word "butcher" and rhymed with "good", and the labels "sut" and "dut" corresponded to the vowels found in the words "subtle" and "Dutch", respectively (a similar process was adopted by Bohn & Bundgaard-Nielsen, 2008).

3.6 Production Tasks

NE speakers produced the English stimuli and fillers in wordlist-reading tasks with two carrier phrases: "He said…and left" (Recording 1) and "The next word is…" (Recording 2), which were visually presented in random order using PowerPoint. In addition to the development of the perceptual tests and training sessions, productions in Recording 1 were used for comparisons with CYG learners' productions; Recording 2 was only used in the elicited imitation task described below. A total of 51 phrases were produced by each participant in each task (37 target words, 10 fillers, 4 practice items at the beginning) and a total of 740 recordings were collected (37 target words x 10 participants x 2 tasks) by NE speakers, but only the target words in Recording 1 were used for analysis, resulting in a total of 307 items (after excluding items based on the ratings of experienced and naïve raters as described in 3.5).

CYG participants produced all stimuli in a wordlist-reading task and in an elicited imitation task. In the wordlist-reading task, the target words along with fillers were embedded in the carrier phrase "He said…and left", which were visually presented to participants in random order using PowerPoint. Participants were asked to produce the wordlists clearly, in normal speaking rate, with the instruction to read and speak as if they were performing a reading task for their classmates in class. In the elicited imitation task, CYG participants listened to the target words and fillers presented in the carrier phrase "The next word is…" (Recording 2 of NE Speaker 4) via headphones, and repeated them in a new carrier phrase

“Now I say…for you” (similarly to Thomson & Derwing, 2016). The recordings of NE Speaker 4 were chosen for the presentation of stimuli in this task, since he had the clearest and most target-like productions based on the ratings of experienced and inexperienced judges. Each task consisted of 51 phrases (37 target words, 10 fillers, 4 practice items at the beginning) and was recorded twice for each participant.

The purpose of the elicited imitation task was to examine whether CYG learners’ productions were affected by English spelling as suggested in previous studies (e.g. Giannakopoulou, Uther, & Ylinen, 2013b, 2017), and if so, to examine their performance in the absence of such orthographic cues. Including two production tasks also enabled an evaluation of task effect, as previous research indicates that the quality of speech produced is affected by task type (e.g. Cucchiarini, Strik, & Boves, 2002; Derwing, Rossiter, Munro, & Thomson, 2004; Detey, Racine, Eychenne, & Kawaguchi, 2014). More specifically, previous meta-analyses reported contradictory findings in relation to production task types; for instance, Lee et al. (2015) found larger effects in studies with tasks involving controlled rather than spontaneous speech, while Hu et al. (2022) reported the opposite pattern. Furthermore, both tasks involve limitations; for instance, reading tasks may be affected by literacy skills (phoneme-to-grapheme mappings), while repetition tasks confound production and perception skills (Nagle & Baese-Berk, 2022), meaning that each of these tasks used in isolation may not fully reveal learners’ true production abilities. At the same time, while other tasks such as picture-naming or spontaneous speech production avoid these confounding effects, they restrict researchers in the choice of items and offer limited control over participants’ productions of the target sounds and the contexts in which they are produced (Nagle & Baese-Berk, 2022). Therefore, in the present study, more controlled tasks were preferred in order to ensure that all target vowels in matching contexts would be elicited.

Finally, CYG speakers produced the CYG stimuli twice in a wordlist-reading task in normal speaking rate. The target words and fillers were embedded in the carrier phrase “Είπεν του…τζι’ έφυε” (/ˈipendu ... ˈdʒefie/, “He said to him…and left”) presented in random order in PowerPoint, using the spelling conventions of the dialect. Participants read 24 phrases in this task: 10 target words, 10 fillers and 4 practice items at the beginning. These productions enabled the acoustic analysis of CYG vowels, which were likely to differ from SMG vowels, as

well as allowing the plotting of vowels in order to visualise their relative positions and measure cross-linguistic similarity/dissimilarity of vowels in the two languages.

Production data were collected using a Zoom H1 audio recorder (sampling rate 44.1 kHz). The data collection process took place at quiet university rooms or labs in the UK and Cyprus. Recordings were saved and analysed as separate sound files representing individual productions of each word. A total of 6216 English target words (37 target words x 14 participants x 2 repetitions x 2 tasks x 3 times) and 280 CYG target words (10 target words x 14 participants x 2 repetitions) were collected by CYG participants. Out of the two repetitions, the best production was chosen for analysis, as determined by the quality of the recording (e.g. noise by participant movements), the quality of the target vowel (most target-like), voice quality (e.g. productions with breathy or creaky voice were avoided), hesitation in producing the target word or mispronunciation of the target word (e.g. errors in reading the target word in wordlist-reading or mis-perception of target word in the elicitation task), resulting in 3108 English items. A further 304 English words produced by CYG learners were excluded due to the same issues, resulting in a total of 2804 target words remaining for analysis. No CYG items needed to be excluded through this process.

3.7 Pre-test

The pre-test was completed by all CYG participants and included both a perceptual and a production task. During the pre-test, participants firstly completed an 11-alternative FCID task without feedback (**Figure 1**), which included 148 natural stimuli produced by 8 of the 10 NE speakers, in only the /bVt/ and /gVt/ contexts. All 11 choices were available to participants, similarly to other studies (e.g. Rato & Carlet, 2020), so that any confusion patterns deviating from the expectations would be identified. Although participants received no immediate feedback on their responses, they were provided with the total number of correct and incorrect responses at the end of the perceptual task. Twenty randomly selected items were added at the beginning as a practice session, so that they familiarise themselves with the procedure. After the perceptual task, CYG learners completed the production tasks (elicited imitation task and wordlist-reading) as described in 3.6. The average time to complete the perceptual task was 19 minutes (range: 12-29 minutes) and the production task was approximately 20 minutes long.

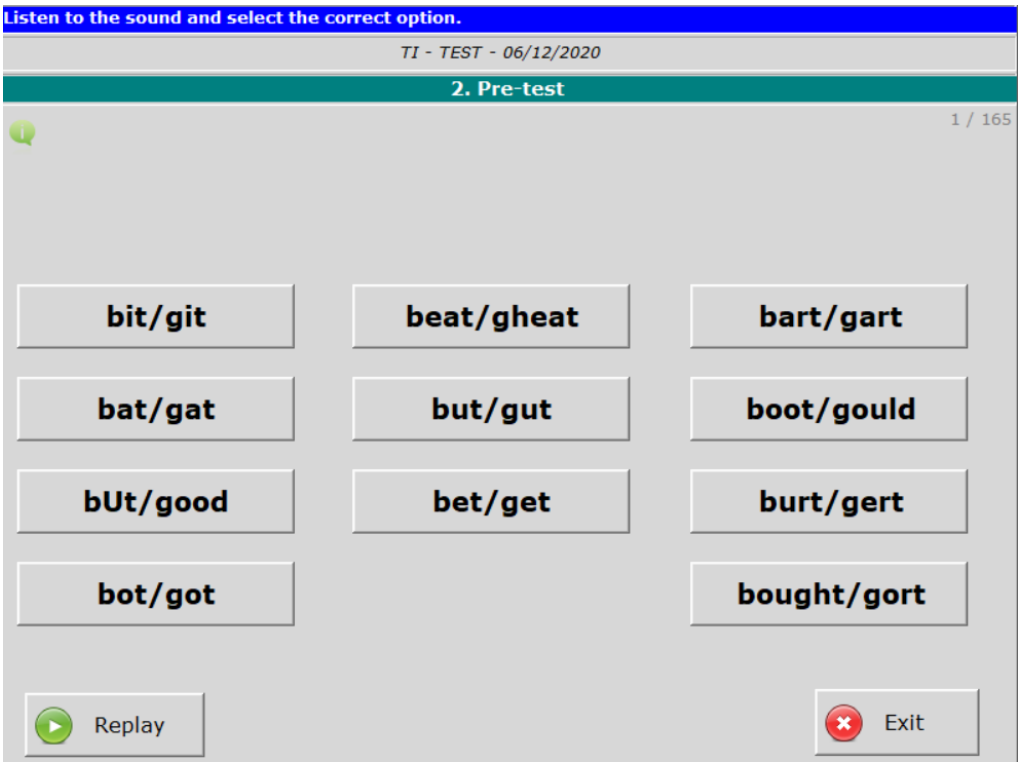


Figure 1. Example of FCID task without feedback

3.8 Training

The training consisted of FCID tasks with immediate feedback. This is the most common paradigm used in HVPT and was found to be superior to other methods such as the categorical discrimination task in training L2 vowel perception and in promoting the generalisation of knowledge to real words (Carlet, 2017, 2019; Cebrian et al., 2019; Lee et al., 2015; Lengeris, 2018; Logan & Pruitt, 1995; Sakai & Moorman, 2018; Thomson, 2018). Providing feedback to trainees has also been found to be more beneficial than training without feedback (e.g. McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002), since it enables learners to direct their attention to acoustic properties of the target segments that are relevant for distinguishing L2 phonemes and ignore irrelevant speaker variability, thereby facilitating L2 category formation (Thomson & Derwing, 2016) by enabling them to use the positive or negative feedback to reinforce existing categories or change them, respectively (Hardison, 2003; Logan et al., 1991). The training included the stimuli elicited from 8 of the 10 NE speakers (Recording 1) and was limited to two phonetic contexts (/bVt/ and /gVt/) so as to avoid confusing and overwhelming the participants, since "unconstrained variability may be counterproductive" (Thomson, 2018, p. 219), while at the same time maintaining some phonetic variation (Thomson, 2011, 2012).

Learners listened to the stimuli at a comfortable volume through headphones and chose their response by clicking on labelled buttons. Trainees were able to listen to each stimulus up to 9 times (maximum repetitions allowed in TP) before selecting their response. If correct, a green tick appeared next to the correct label, and they could proceed to the next stimulus (**Figure 2**). The message "Incorrect answer! Click on Replay before checking the answer" and the correct response appeared when an incorrect label was chosen, and participants had to listen to the stimulus again before proceeding (**Figure 3**). The total number of correct and incorrect responses was also provided at the end of each session.

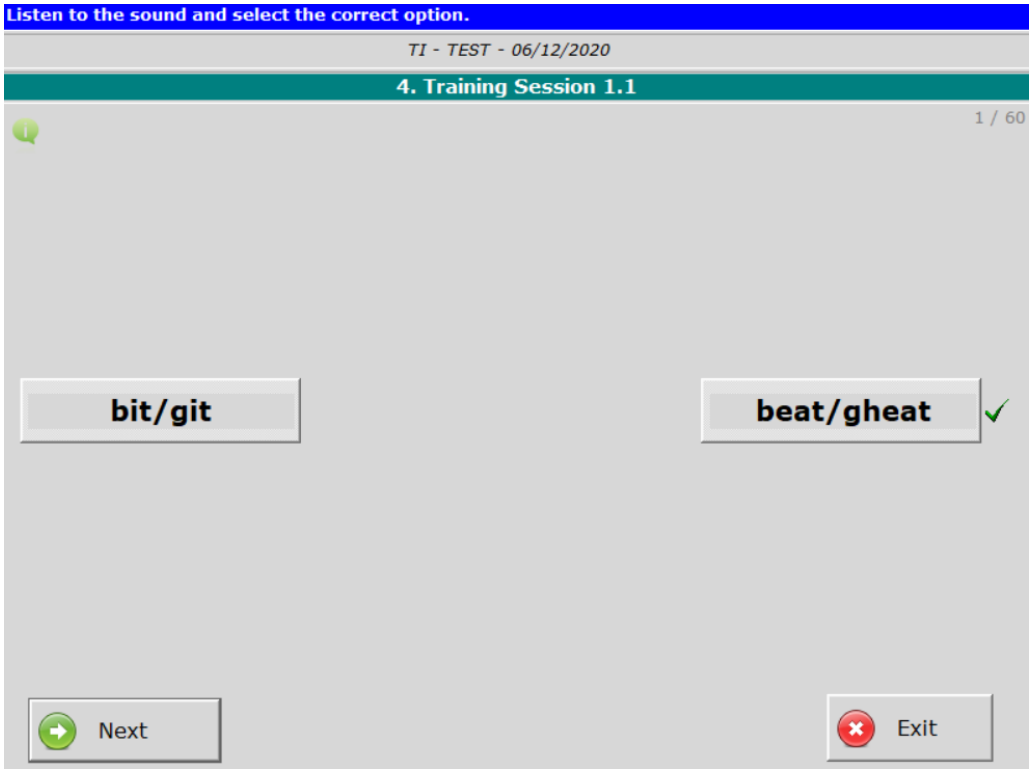


Figure 2. Example of FCID task with feedback (correct response)

Trainees completed eight training sessions over three weeks, on a flexible schedule, and at their own place and time. A maximum of one session per day was completed and the researcher observed the completion of each session via online screen-sharing and remote desktop software. Each training session was blocked by speaker (as in previous studies, e.g. Lengeris & Hazan, 2010; see also Thomson, 2018) and contained 330 trials. Blocking sessions

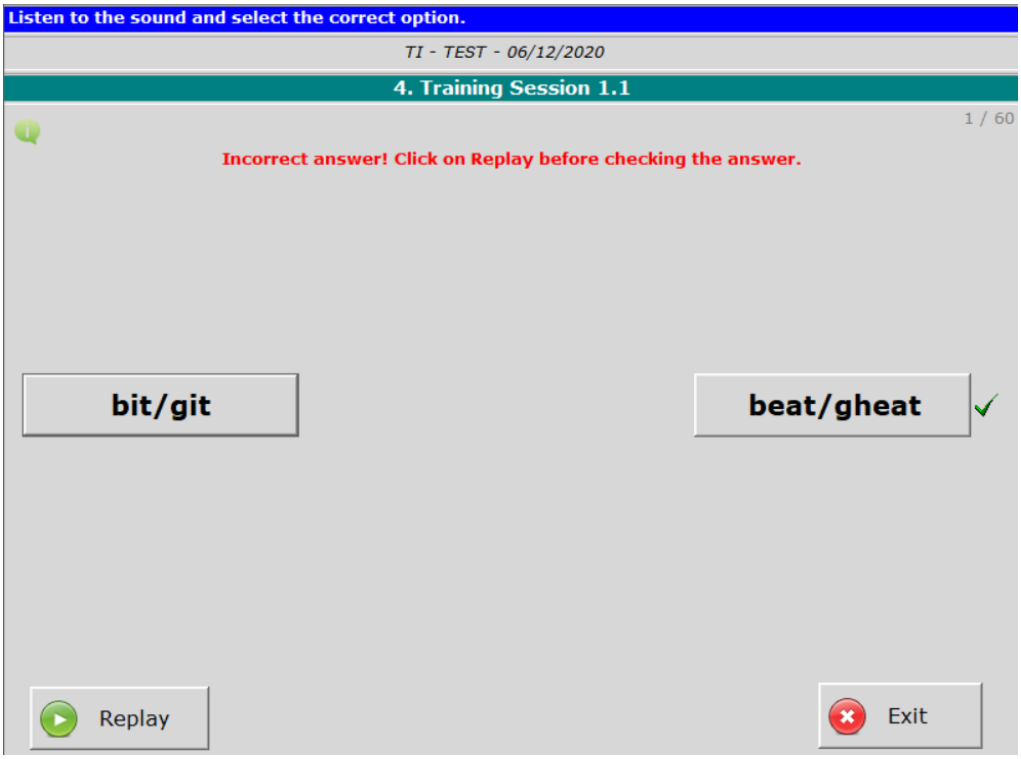


Figure 3. Example of FCID task with feedback (incorrect response)

by speaker was preferred, given that speech processing becomes more difficult when the listener has to adapt to a different speaker on each trial, which is even more problematic for non-native speakers (Brekelmans et al., 2022). The average time to complete each training session was 32 minutes (range: 17 minutes to 1 hour). Participants were not prompted to produce the target vowels in any way during the training.

The whole set of target vowels in the /bVt/ and /gVt/ contexts was included in each session. This was based on Nishi and Kewley-Port (2007), who demonstrated that training on the full set of vowels instead of a smaller subset of difficult vowel categories resulted in larger perceptual gains, which were also extended to new speakers and words and were maintained three months later. However, each session was divided into five training blocks, each focusing on only one vowel contrast (/ɪ-i:/, /e-ɜ:/, /ɑ:-æ-ʌ/, /ɒ-ɔ:/ and /ʊ-u:/), similarly to Wang and Munro (2004). This grouping of vowels was determined by the perceptual mapping of L2 English vowels onto CYG vowels as observed in previous studies, and correspond to the CYG vowels /i, e, a, o, u/ respectively (Georgiou, 2019; see 2.5.3). A contrastive set rather than contrastive pairs were preferred for the /ɑ:-æ-ʌ/ contrast, to ensure that learners received an equal amount of stimuli for each target vowel, since splitting them into separate pairs would mean providing additional stimuli for this than for other contrasts. Although this made

identifying the members of this contrast more difficult for learners, it was considered appropriate since it resembles the difficulty that can emerge in naturalistic speech as well. The blocks were presented in random order in each session, but the order was the same for all participants.

Upon hearing the stimulus, trainees clicked on the labelled button containing the word they heard, choosing from a set of two or three options, which were set as follows: “bit/git” or “beat/gheat” for the /ɪ - i:/ contrast, “bet/get” or “burt/gert” for the /e - ɜ:/ contrast, “bot/got” or “bought/gort” for the /ɒ - ɔ:/ contrast, “bUt/good” or “boot/gould” for the /ʊ - u:/ contrast, and “bart/gart”, “bat/gat”, or “but/gut” for the /ɑ: - æ - ʌ/ contrast. Providing a limited set of options for the training was preferred, as in this way, participants would be forced to focus on distinguishing between the members of problematic contrasts, instead of contrasts that might not pose any problems for these learners. Participants could have a break after each block.

3.9 Immediate Post-Test and Generalisation Test

Similarly to other studies (e.g. Carlet & Cebrian, 2014; Cebrian et al., 2019; Hazan et al., 2005; Hutchinson & Dmitrieva, 2022; Wang & Munro, 2004; Yuan & Archibald, 2022), the post-test used to determine whether there was any improvement in performance was identical to the pre-test, with the addition of a generalisation test, to evaluate whether any knowledge that CYG participants acquired from the training was extended to untrained contexts and speakers. The generalisation test consisted of 3 blocks: the first included 36 known stimuli produced by the 2 new speakers (New Speakers test), the second included 98 stimuli in new contexts produced by familiar speakers (New Contexts test), and the third had 25 stimuli in new contexts produced by new speakers (New Speakers and Contexts test). The stimuli in each block were randomised and included all target vowels. Each new word with its corresponding label was presented before the test, and participants completed a 20-item practice session at the beginning of the generalisation test, so that they were introduced to the new labels, which are shown in **Figure 4**. In the production task, CYG participants only produced the English stimuli in the same wordlist-reading and elicited imitation tasks described in section 3.6.

The stimuli in both the production and perception tasks were randomised in the post-test, to limit familiarity effects. Possible effects of familiarity with the tasks and stimuli were also addressed by the inclusion of the control group (Carlet & Cebrian, 2014). The average

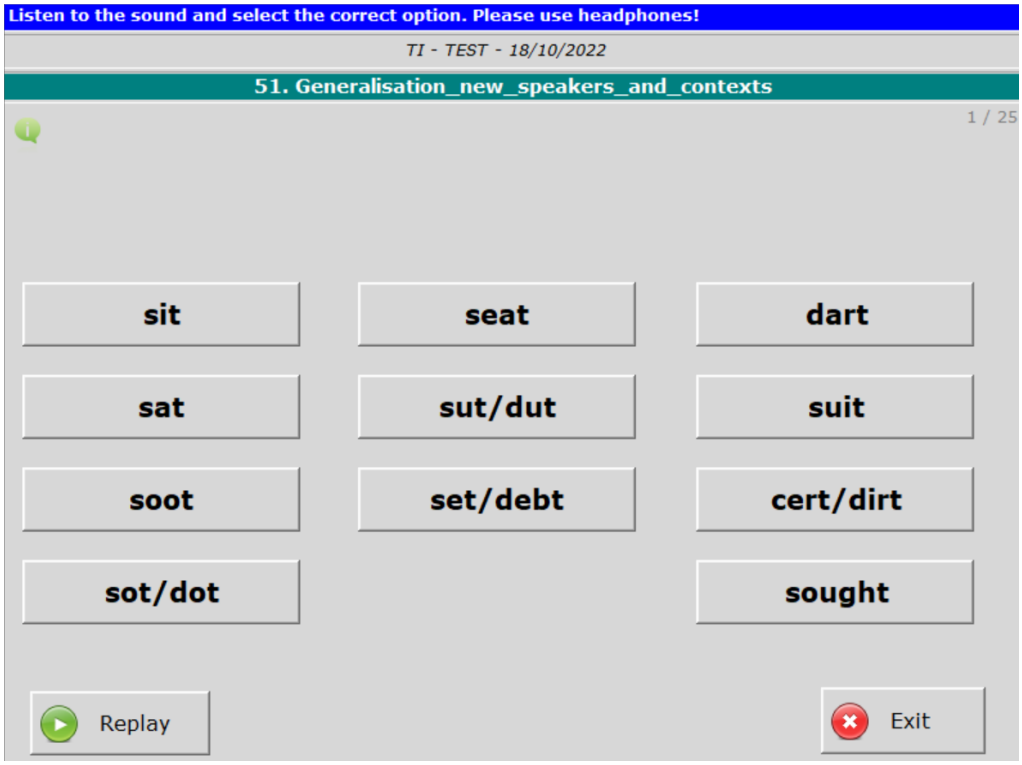


Figure 4. Example of generalisation test

time to complete the perceptual tasks was 26 minutes (range: 16-55 minutes) and the production task was approximately 20 minutes long. Participants could have a break between the perceptual and production tasks, as well as halfway through the stimuli in each task.

Finally, a post-training evaluation questionnaire was completed by trainees, in order to assess the content and construct validity of the training (Sandelowski, 2000). Since the aim of this thesis was not to evaluate the different methodological approaches used in HVPT, the data obtained through this questionnaire have not been analysed further. However, this is important in taking into consideration participants' comments in the development of training paradigms in future studies, and to get a sense of participants' state of mind during the training, i.e. whether they were tired, frustrated etc.

3.10 Retention Test

A retention test was also conducted to obtain information on the long-term effects of the training. The retention test was identical to the immediate post-test including the generalisation test with stimuli randomised in all tasks, and it was used to investigate whether the effects of the training persisted two months later. The average time to complete the perceptual task was 26 minutes (range: 17-58 minutes) and the production task was

approximately 20 minutes long. Participants could have a break between the perceptual and production tasks, as well as halfway through the stimuli in each task.

3.11 Quantitative Analysis

Perceptual performance was measured through %-correct identification scores of participants and patterns of confusion. Statistical analyses to examine the effect of the IVs on the dependent variable (DV), i.e. participants' performance, were conducted using mixed-effects binomial logistic regression in R (R Core Team, 2022) v 4.2.1, since participants' responses were either correct or incorrect and random effects were needed to account for inter-subject variance. The general formula used in these analyses was *glmer (DV~IV+(1|Subject)+(1|Vowel), data=data.frame, family="binomial")*. The significance level was set at 0.05.

Production data were analysed both acoustically and through intelligibility ratings by NE listeners. Although acoustic analyses are a relatively objective measure, they still involve methodological choices that may affect the results (Nagle & Baese-Berk, 2022). As Thomson (2018, p. 225) noted, acoustic analyses can provide insights into changes in the productions of learners, but "only human raters can accurately determine to what extent that change is on target, and whether it contributes to the speakers' intelligibility…for listeners". At the same time, while listener-based measures offer broad ecological validity, listener perceptions may be biased as an effect of various factors such as lexical frequency or social information (Nagle & Baese-Berk, 2022). Furthermore, Baker and Trofimovich (2005, 2006) pointed out that although human raters (whether trained or untrained) can help us determine larger-scale differences in production (e.g. whether learners produce two vowels differently), they are less likely to identify subtler differences, as opposed to acoustic analyses which provide a more fine-grained measure of production accuracy. Therefore, using a combination of tasks to measure production performance similarly to some previous studies (e.g. Lambacher et al., 2005; Lengeris & Hazan, 2010; Lengeris, 2009a, 2018) was considered a more appropriate approach to obtain more comprehensive results and minimise the limitations that each of them entails. The two methods are described in 3.11.1 and 3.11.2.

Finally, in order to assess the role that language use patterns, motivation and input play on learners' perceptual and production performance (RQ 5), a correlational design was used to assess the relationship between these variables, measured through the relevant

questionnaire items, and the %-correct identification scores and intelligibility scores of CYG learners for perception and production, respectively.

3.11.1 Acoustic Analyses

The productions of NE speakers and CYG learners were acoustically analysed using Praat to identify similarities and differences. Similarly to previous studies (e.g. Lengeris, 2009a), the target vowels were segmented manually and duration and formant values were measured using simultaneous inspections of the waveform and spectrogram. Vowel onset and offset were identified based on the onset and offset of periodic energy in F2 and higher formants in the spectrogram as shown in **Figure 5** (Nishi, Strange, Akahane-Yamada, Kubo, & Trent-Brown, 2008; Themistocleous, 2017b). Vowel data were extracted using the script in Appendix C.2 although formant settings were adjusted for individual speakers to address the fact that longer vocal tracts have lower formant frequencies and vice versa. Only F1 and F2 measurements were included, since these are the most significant formant frequencies for the classification of vowels (Themistocleous, 2017b). The F1 (high-low dimension) and F2 (front-back dimension) determine the position of a vowel in the vowel space, and are particularly important, as when the two formants are plotted, vowels can be represented within their vowel space (Themistocleous & Logotheti, 2016). F1 and F2 values were measured at mid-point (Hutchinson & Dmitrieva, 2022), where vowels exhibit the least effect from neighbouring segments. Outliers in duration and formant values were identified and removed before proceeding with any statistical analyses.

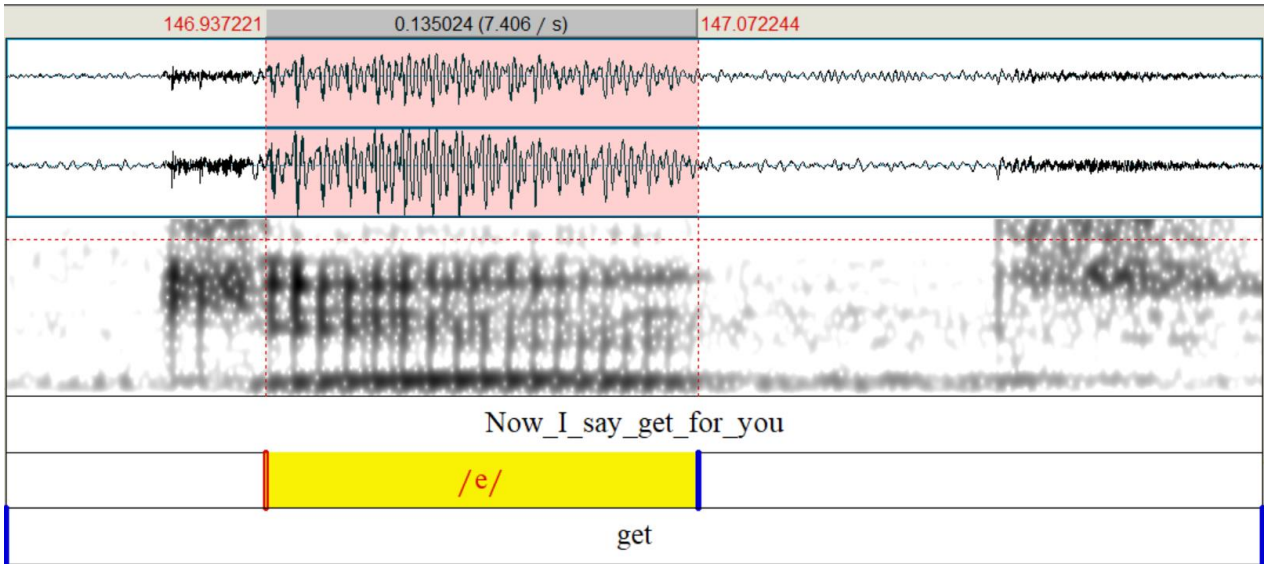


Figure 5. Example of segmentation

Effects of physiological differences between male and female participants were removed through normalisation using the Lobanov method, as it was found to be among the most effective techniques by Clopper (2009) and Flynn (2011) among many others. NORM (Thomas & Kendall, 2015a) v 1.1 was used for normalisation, and data were normalised as follows: all CYG-participant productions of L1 and L2 vowels at all three times were normalised as one set, and all NE-participant productions of the target vowels were normalised as another set. It was not considered appropriate to normalise the values of the two groups (NE and CYG) together, given the disadvantage of vowel-extrinsic methods (including Lobanov) in dealing with comparisons between different languages that contain different vowel systems. In this case, the fronting of GOOSE in the productions of NE speakers would cause "the whole vowel system to be weighted toward front vowels" (Thomas & Kendall, 2015b). Therefore, since CYG learners' productions lack this fronting, it would not be appropriate to normalise the values of the two groups together.

A 10% representative sample of the recordings was submitted to intra-rater reliability assessment, to ensure that the acoustic measurements and annotations were correct (Mildner, 2013). The sample included 315 target stimuli: 280 from CYG learners' productions across group, context, time and task and 35 for NE speakers from Recording 1. Absolute differences in duration between first and second measurement did not exceed 10ms in any case, and a paired-samples t-test conducted in R showed no significant difference between the two measurements ($p=0.143$). The Pillai score was also calculated to assess differences in spectral characteristics between the original and sample measurements in the un-normalised productions of CYG learners only, since the sample representing NE speakers was very small. The Pillai score was calculated for each target vowel separately in R, using the formula *manova (cbind (F1, F2)~Measurement, data=data.frame)*. The highest Pillai score obtained was 0.004, suggesting a very high degree of overlap between the first and second measurements in the sample.

Statistical analyses on duration were conducted in R, using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). Linear mixed-effects models enable the analysis of data with unbalanced designs and with missing values, as well as taking into account variability within and across participants (Brown, 2021). As in previous studies (e.g. Nimz & Khattab, 2020) separate models were conducted for each analysis, where *subject* was modelled as a random

factor and *context*, *task*, *group*, *subgroup*, *vowel* and/or *time* as fixed factors, as indicated in each analysis. The general formula used was: *lmer (DV~IV1*IV2 + (1|Subject), data=data.frame)*. The significance level was set at 0.05. Vowel overlap was measured by calculating the Pillai score in R, which was firstly used by Hay, Warren, and Drager (2006) and was found to be superior to other methods in Kelley and Tucker (2020). The general formula used in R was: *manova (cbind (F1, F2)~IV, data=data.frame)*. Higher Pillai scores indicate a smaller amount of vowel overlap whereas lower Pillai scores indicate a higher amount of overlap between two vowels (Hay et al., 2006). The data frames included individual target vowels or vowel pairs as stated in each part of the analysis, in order to facilitate the interpretation and reporting of results, particularly given the fact that participants were likely to perform differently in different vowels or vowel pairs across tasks and times. The target vowels as produced by NE speakers and CYG learners were also plotted in the vowel space to visualise their position and enable comparisons.

3.11.2 Intelligibility Measurements

Intelligibility ratings were included to complement acoustic analyses, in order to assess how L2 productions are evaluated by NE listeners and determine whether any adjustments made by the learners were perceptually salient, i.e. apparent to native listeners (Hutchinson & Dmitrieva, 2022). Furthermore, as Nagle and Baese-Berk (2022) noted, perception and production measures should be paired in a comparable way in order to evaluate the perception-production link in a valid way. This was achieved by the use of this categorical production measure, which mirrored the perceptual tasks completed by learners. Intelligibility instead of accentedness or comprehensibility measurements were preferred, since the latter two are considered to be subjective measures, whereas intelligibility constitutes a more objective method to test what listeners actually understand, without prejudice or beliefs that might influence their responses (e.g. willingness to make an effort to understand) (Baese-Berk et al., 2020; Hu et al., 2022).

Previous research varies as to the recruitment of experienced or naïve listeners for the intelligibility tasks, as well as the number of raters included. For example, Thomson and Derwing (2016) recruited two phonetically-trained raters, Lengeris (2009a, 2018) and Lengeris and Hazan (2010) included two native raters with unspecified experience, Georgiou (2021) included three naïve raters, Suzukida and Saito (2019) recruited four experienced EFL/ESL

teachers with extensive phonological knowledge and familiarity with the accented speech of their learners (for comprehensibility ratings), Carlet (2017) included 12 English teachers, Lambacher et al. (2005) recruited 26 phonetically-trained judges, whereas Derwing and Munro (1997) recruited 26 naïve listeners. The length of rating scales also varies in previous studies; for example, Thomson and Derwing (2016) included a 3-point scale, Bohn and Steinlen (2003) and Birdsong (2007) used a 5-point scale, whereas Lengeris and Hazan (2010) had a 7-point scale and Suzukida and Saito (2019), Lengeris (2009a, 2018) and Carlet (2017) included a 9-point scale.

Isaacs and Thomson (2013) examined whether rater experience (experienced vs. novice raters) and scale length (5- or 9-point scale) affected raters' judgments of L2 speech in terms of comprehensibility, accentedness and fluency, and found that neither factor yielded significant differences between the groups. This was subsequently supported by Hu et al.'s (2022) meta-analytic review, which showed that "the length of the scale did not moderate the effectiveness of interventions" (p. 16). Furthermore, Derwing and Munro (1997) and Derwing et al. (2004) suggest that untrained raters can be reliable in assessing comprehensibility, accentedness and fluency of non-native speech. Therefore, since the purpose of the current study was to examine how the speech of L2 learners is evaluated in naturalistic conditions that resemble real-life situations, naïve listeners were preferred.

For this task, 5 naïve NE listeners provided intelligibility measurements through an 11-alternative FCID task without feedback, including a goodness-of-fit rating of the productions of CYG speakers on a scale from 1 (very bad) to 5 (excellent) administered using TP. A longer rating scale was avoided, since raters might have found it difficult to differentiate between the steps in a 7- or 9-point scale (Isaacs & Thomson, 2013), particularly since they were not experienced phoneticians. The 5 raters were chosen as to form as homogeneous a group as possible both among them and with the NE and CYG groups that completed the previous tasks. They were native speakers of SSBE aged between 18-28 years old, and they reported that they did not speak any other languages or have a speech or hearing impairment. These raters received financial compensation for their participation. Before proceeding with the ratings, raters were asked to complete a 37-item familiarisation task, which included the target words produced by 2 NE speakers (1 male, 1 female) and had to be completed without errors in order for raters to proceed. In addition to familiarising themselves with the labels and procedure,

this task also served as a tool to confirm that raters were SSBE speakers and able to accurately identify all SSBE target vowels.

Similarly to Lengeris (2018), it was preferred to present the target vowels in their context rather than presenting isolated vowels, since the latter does not resemble naturalistic conditions. Using single words in the intelligibility task was not considered problematic, since this study focuses on pronunciation rather than grammar or fluency (similarly to Suzukida & Saito, 2019), but also because Munro, Derwing, and Burgess (2010) have demonstrated native listeners' sensitivity in identifying non-native speech even when provided with a single word. At the same time, as noted by Melnik-Leroy et al. (2022), target vowels in isolation might have been too short for raters to assess, while larger segments could have yielded biased responses influenced by the speaker's global accent, even if raters were instructed to focus on specific sounds. Even though rating full words may entail a lexical bias (Melnik-Leroy et al., 2022), the inclusion of non-words in the present study mitigated this limitation, rendering this the most preferrable option for stimuli presentation. However, as in Flege and Wayland (2019), the stimuli were edited so that consonant production would not influence listeners' ratings of the target vowels. More specifically, word-initial voiced stops were produced with varying amounts of pre-voicing, which was digitally removed. Furthermore, any instances of post-vocalic /r/ or /l/ were also removed when present in the productions of learners. In such cases, portions of the same participant's productions of word-final /t/ or /d/ in other stimuli were used to replace these consonants if affected by the preceding approximant. Upon these modifications, all stimuli were checked for naturalness before administering the task.

Listeners were instructed to listen to the stimuli carefully through headphones and click on the labelled button corresponding to the word they heard, choosing their response from the whole set of 11 options. Raters were able to listen to each stimulus up to 9 times before selecting their response. After making their choice, listeners had to indicate whether the word they heard was a good or bad example of the word in the label using the scale. The procedure followed was similar to Iverson et al. (2012): each listener rated all 2804 target words produced by learners in all tasks and times, and these were subdivided into 10 tasks, each containing an average of 285 stimuli (range: 270-302). Each task contained randomly selected stimuli across participants, production tasks and times, blocked by context and presented in random order. Raters were instructed not to complete more than one task per day, in order

to avoid the effects or raters' fatigue, which can affect the accuracy of ratings (Suzukida & Saito, 2019). This was confirmed through the excel file automatically generated by TP which records participants responses as well as the date of completion and time spent on the task. Furthermore, participants were instructed to use the scale flexibly.

Inter-rater reliability was calculated using the Cronbach's alpha (Isaacs & Thomson, 2013; Saito et al., 2022; Saito, Trofimovich, & Isaacs, 2017; Suzukida & Saito, 2019) to assess raters' internal consistency. The Cronbach's alpha for raters' correct or incorrect identification of the stimuli was 0.82 showing high agreement among the raters. However, raters' responses on the quality of the stimuli on the 5-point rating scale had a very low internal consistency (α=0.42), and therefore were not included in further analyses. Raters' identification scores were analysed using a mixed-effects binomial logistic regression in R using the general formula *glmer (Response~IV+(1|Vowel)+(1|Rater), data=data.frame, family="binomial")*, where the IV was adjusted as indicated in each analysis.

3.12 Qualitative Analysis

Following an explanatory sequential mixed methods design (QUAN→qual), the quantitative data analysed were used to purposefully select individuals for the qualitative phase (Creswell, 2014). Given the limitations of questionnaires in measuring the variables under examination (Flege, 2008), qualitative data were collected using criterion sampling and representing extreme cases (Sandelowski, 2000). More specifically, three CYG learners participated in a follow-up semi-structured interview (see sample questions in Appendix B.4), which complemented the questionnaire data, in order to gain additional insights (Schleef, 2014). A similar procedure was followed by Tsang (2022). The interviewees were trainees identified based on the results of the intelligibility task completed by NE raters, representing the highest and lowest performance. Interviews were conducted in the language of their preference (Greek or English), and at a quiet place of their preference in person or through Microsoft Teams. The interviews, which lasted approximately 30 minutes, were recorded and transcribed through Microsoft Teams upon the consent of participants, who were also informed that they could withdraw at any point and could choose not to answer any questions that made them feel uncomfortable.

Qualitative data were analysed using NVivo (QSR International Pty Ltd, 1999-2023) v 14, and following the six-phase thematic analysis by Braun and Clarke (2006) in a deductive way:

1. Familiarising with the data, 2. Generating preliminary codes, 3. Searching for themes, 4. Reviewing themes, 5. Defining themes, and 6. Selecting examples for report. The analysis focused on identifying themes relating to motivation, language use patterns and input received.

3.13 Summary

This chapter has presented the RQs and methods followed to address them (3.1) and provided a detailed description of both NE and CYG participants (3.2) as well as the procedures followed and the HVPT protocol and tools used for the purposes of this study (3.3-3.10). Finally, the chapter described the quantitative and qualitative analyses conducted for each set of data collected (3.11-3.12). Some methodological differences between previous studies and the present research have also been pointed out, and justifications for methodological choices have been provided.

The following chapter presents the results of the study as well as details about each analysis conducted. The chapter begins with the findings in relation to learners' perceptual performance and then presents the findings of acoustic analyses on duration and spectral characteristics, which are complemented by the results of the intelligibility ratings of NE speakers. Finally, the results of correlation and qualitative analyses are presented, which provide further insights into individual differences in learners' perceptual and production performance. It should be noted that for ease of reference and in an attempt to make the results of this study comparable to other studies that have chosen a different set of symbols to represent the target vowels, keywords from Wells' (1982) vowel chart will be used instead of phonetic symbols to represent the target L2 vowels of the present research in the remainder of this thesis, as follows: KIT for /ɪ/, FLEECE for /iː/, DRESS for /e/, NURSE for /ɜː/, BATH for /ɑː/, TRAP for /æ/, STRUT for /ʌ/, LOT for /ɒ/, NORTH for /ɔː/, FOOT for /ʊ/ and GOOSE for /uː/.

CHAPTER 4: RESULTS

4.1 Perception

4.1.1 Group Performance in Trained Contexts and Voices

The %-correct identification scores of CYG participants in the perceptual tests at pre-test (T1), post-test (T2) and retention test (T3) in the /bVt/ and /gVt/ contexts are shown in **Figure 6**.

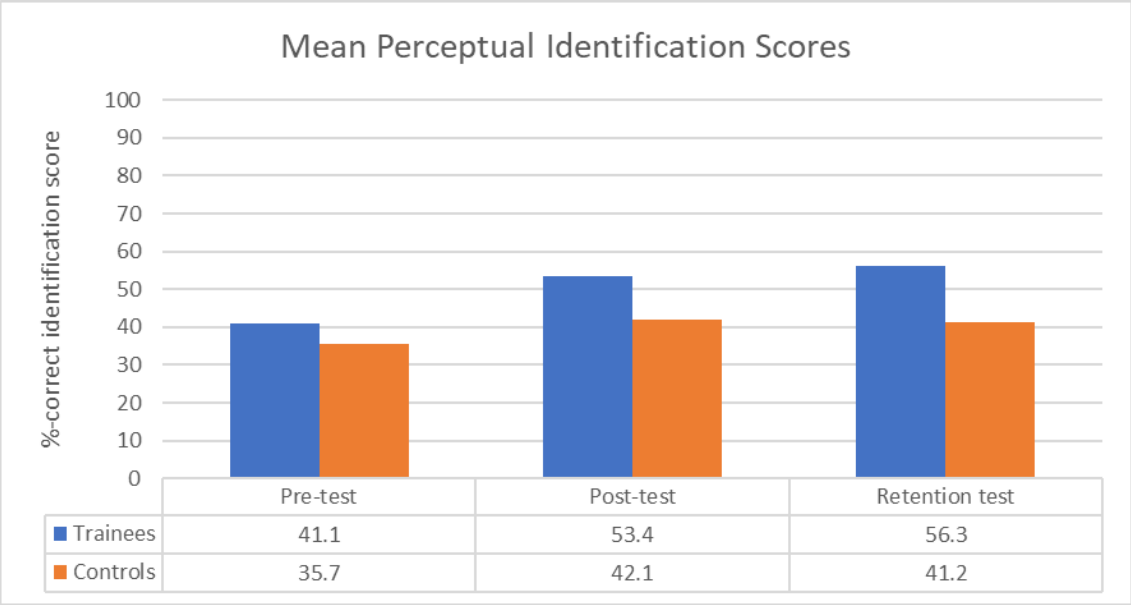


Figure 6. Mean overall perceptual identification scores of trainees and controls in trained contexts at T1, T2 and T3

Trainees performed better than controls at all times, and their performance was improved across time. The performance of controls was also improved at T2, but not at T3. These differences between groups at T1, T2 and T3 were analysed using a mixed-effects binomial logistic regression (formula: *glmer (Response~Time+(1|Subject)+(1|Vowel), data=data.frame, family= "binomial")*). The analysis showed significant differences between the two groups at T3 only (T1: est.=0.23, p=0.188; T2: est.=0.418, p=0.255; T3: est.=0.62, p=0.013). Within-group differences were also assessed using logistic regression, to examine each group's performance across time. In both groups, differences in perceptual identification reached significance between T1 and T2 (trainees: est.=0.527, p<0.001; controls: est.=0.271, p=0.006) and between T1 and T3 (trainees: est.=0.64, p<0.001; controls: est.=0.236, p=0.016). Differences between T2 and T3 did not reach significance for either group.

4.1.2 Individual Participant Performance in Trained Contexts and Voices

Each participant's individual performance as well as changes between the three time points are shown in **Table 3**. The analysis of the performance of each participant across the three time points using binomial logistic regression (formula: *glmer (Response~Time+(1|Vowel), data=data.frame, family= "binomial")*) showed significant differences between T1 and both T2 and T3 for participants CYG03 (T2: est.=0.653, p=0.009; T3: est.=0.934, p<0.001), CYG04 (T2: est.=0.978, p<0.001; T3: est.=1.156, p<0.001), CYG08 (T2: est.=1.21, p<0.001; T3: est.=1.148, p<0.001), CYG12 (T2: est.=0.813, p=0.002; T3: est.=0.567, p=0.025), CYG15 (T2: est.=0.631, p=0.011; T3: est.=0.49, p=0.049) and CYG17 (T2: est.=1.298, p<0.001; T3: est.=0.866, p<0.001). Significant differences between T1 and T2 were observed in participant CYG16 (est.=-1.373, p<0.001), while differences between T1 and T3 were observed in participants CYG01 (est.=0.506, p=0.039) and CYG06 (est.=0.748, p=0.003). These results show that most but not all trainees improved in their identification scores across time, while the performance of CYG16 was lower at post-test compared to the pre-test. At the same time, two controls (CYG06 and CYG15) also demonstrated improvement across time.

SUBJECT	T1	T2	T3	GAIN (T1-T2)	GAIN (T1-T3)	GAIN (T2-T3)
CYG01*	51.4	58.8	62.8	7.4	11.5	4.1
CYG03*	39.2	53.4	59.5	14.2	20.3	6.1
CYG04*	45.3	65.5	68.9	20.3	23.7	3.4
CYG05*	39.9	40.5	50	0.7	10.1	9.5
CYG06	35.8	44.6	52	8.8	16.2	7.4
CYG07	33.8	42.6	39.9	8.8	6.1	-2.7
CYG08*	38.5	67.6	66.2	29.1	27.7	-1.3
CYG09	43.9	45.3	47.3	1.4	3.4	2.0
CYG10	29.7	39.2	30.4	9.5	0.7	-8.8
CYG11	42.6	38.5	38.5	-4.1	-4.1	0.0
CYG12*	55.4	72.3	67.6	16.9	12.2	-4.7
CYG15	28.4	42.6	39.2	14.2	10.8	-3.4
CYG16*	28.4	10.1	26.4	-18.2	-2.0	16.2
CYG17*	31.1	58.8	49.3	27.7	18.2	-9.5

Table 3. Individual participant performance at T1, T2 and T3, and numerical changes between time points. Asterisks indicate trainees.

4.1.3 Performance per Vowel in Trained Contexts and Voices

Figure 7 shows CYG learners' mean perceptual performance in each target vowel at T1, and **Table 4** shows learners' mean %-correct identification scores of each vowel in each context. Binomial logistic regressions with *Vowel* as the fixed and *Subject* as the random effects (*glmer (Response~Vowel+(1|Subject), data=data.frame, family= "binomial")*) were conducted to examine differences in the identification scores between the vowels in a pair at T1 for both groups of learners. The results showed differences between contrastive vowels in almost all pairs (DRESS-NURSE: est.=-1.147, p<0.001; BATH-TRAP: est.=1.24, p<0.001; BATH-STRUT: est.=0.726, p=0.005; TRAP-STRUT: est.=0.455, p=0.031; LOT-NORTH: est.=-0.51, p=0.012), with the exception of KIT-FLEECE and FOOT-GOOSE, where the two vowels were identified at a similar rate.

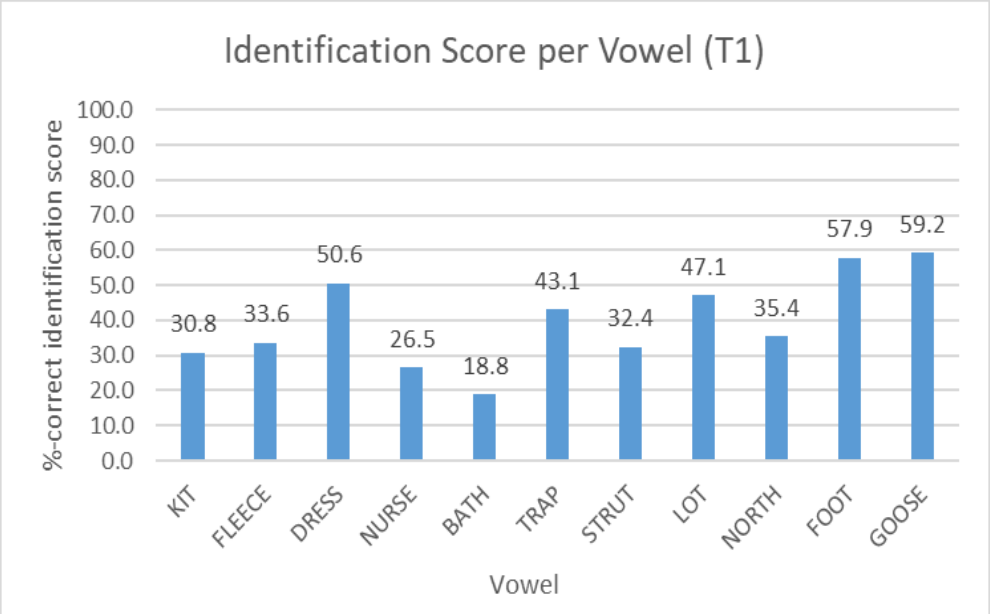


Figure 7. Mean perceptual identification scores of both CYG groups at T1 for each target vowel

Table 5 is a confusion matrix, showing participants' percentage of responses to each target vowel and the most commonly confused vowels at T1, demonstrating how CYG learners perceived L2 vowels before any intervention. Based on this, KIT was most commonly identified as DRESS or itself, with the expected FLEECE being a less common response. FLEECE was most commonly identified as KIT or itself, as expected. DRESS and NURSE were most commonly identified as the correct vowel, with STRUT being the second option at a considerably lower frequency. BATH was most commonly identified as STRUT, with the second response being the

VOWEL	CONTEXT	
	/bVt/	*/gVt/*
KIT	49.1	12.5
FLEECE	36.9	30.4
DRESS	28.6	63.3
NURSE	40.5	16.1
BATH	9.5	30
TRAP	41.4	44.9
STRUT	45.9	16.7
LOT	25.5	66.1
NORTH	50.9	19.8
FOOT	17.9	67.9
GOOSE	62.5	51.1

Table 4. Mean %-correct identification scores of both CYG groups at T1 for each target vowel in each context

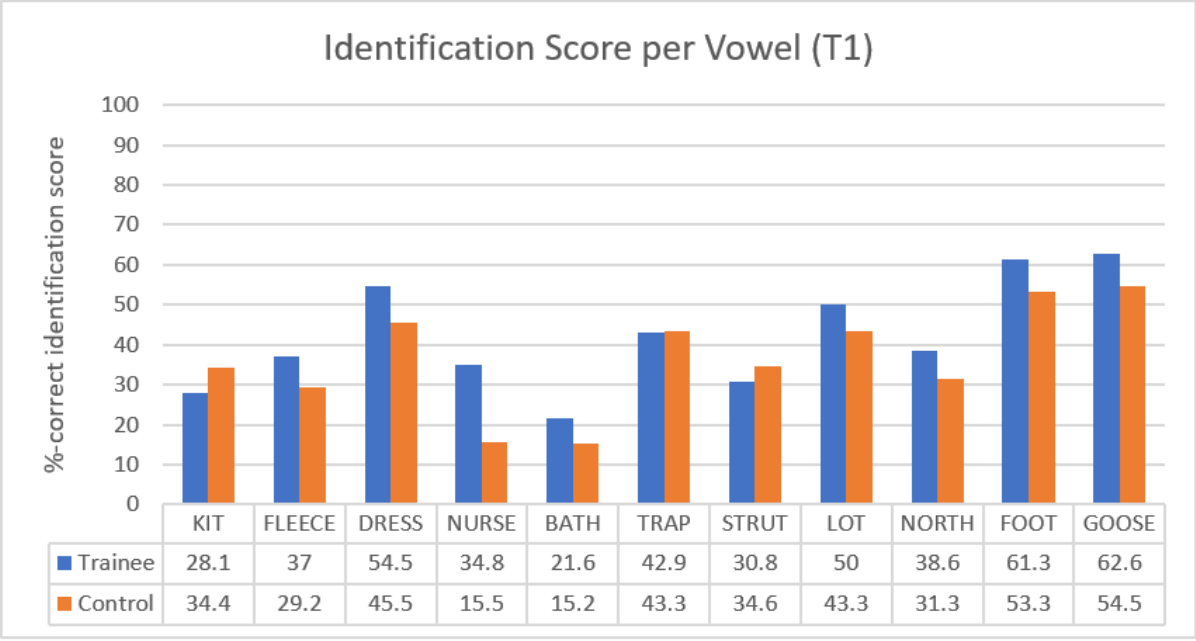
correct vowel and the expected TRAP being the third most common response. TRAP and STRUT were most commonly identified as the correct vowel, with the second response being the other vowel in the pair and the third option being BATH with a lower frequency in both cases. The pairs LOT-NORTH and FOOT-GOOSE were also most commonly identified as the correct vowel, with the second most frequent response being the other vowel in the pair. However, in the case of FOOT-GOOSE, the correct vowel was chosen at a much higher rate, with the second option being considerably less common.

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**30.8**	14.3	**37.5**	5.4		2.7	1.8	1.8		2.7	2.2
FLEECE	**53.4**	**33.6**	6	2.7						1.8	1.3
DRESS	2.6		**36.3**	9.7	5.2	11.7	**13.6**	2.6		2.6	
NURSE		2.6	7	**26.5**	14.8	12.2	**16.8**	8.2	1.5	5.1	4.6
BATH				9.7	**18.8**	16.2	**24.7**	16.2	7.8	3.2	1.9
TRAP				1.9	13.4	**43.1**	**30.1**	4.3	1	2.9	1.4
STRUT	1.1	1.1		6	14.8	**27.5**	**32.4**	11	2.2	3.3	
LOT	1			1.9	5.2	2.9	6.2	**47.1**	**30**	3.3	1.9
NORTH				3.1			1.3	**21.1**	**35.4**	16.6	19.7
FOOT	1.4		5.6	4.3	2.1	2.1	7.1	5	2.1	**57.9**	8.6
GOOSE	4.5	7	1.9			1.3			5.7	**17.8**	**59.2**

Table 5. Confusion matrix of participants' percentage of responses to each target vowel at T1. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

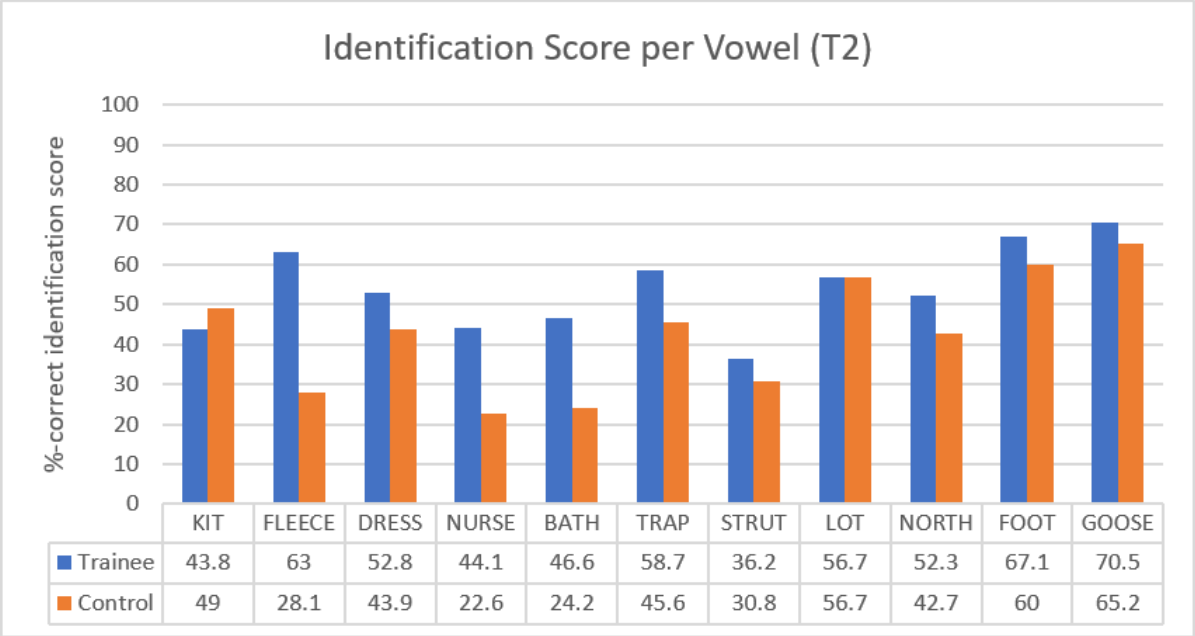
Figures 8-10 show the mean identification scores of trainees and controls in each target vowel at T1, T2 and T3, respectively. In terms of performance in each vowel, it is worth noting that GOOSE and FOOT were better identified than any other vowel at all times by both groups. These were followed by DRESS, LOT and TRAP for both groups at T1. At T2, GOOSE and FOOT were followed by FLEECE and TRAP for trainees, with scores above 50% in LOT, DRESS and NORTH as well; for controls, GOOSE and FOOT were followed by LOT, with scores below 50% for all other vowels. At T3, trainees performed well in most vowels, with scores over 50% in GOOSE, NORTH, DRESS, FLEECE, FOOT, TRAP, LOT and NURSE in this order; controls on the other hand, did not have the same success, with scores at 50% or higher only for GOOSE, FOOT, KIT and LOT. The most difficulties arose in BATH and NURSE, which were in all cases among the four worst-identified vowels. STRUT was also among the most problematic vowels, especially for trainees.

Overall, trainees' performance improved across time for all vowels, except in DRESS, where they performed better at T3 but slightly worse at T2 compared to T1. Trainees' performance was best at T3 for most vowels. Controls also performed better at T2 and/or T3 for most vowels (KIT, NURSE, BATH, TRAP, LOT, NORTH, FOOT, GOOSE), although the improvement in each of these vowels was much smaller compared to trainees. In FLEECE, DRESS and STRUT, their performance remained similar or worse at T2 and/or T3 compared to T1.



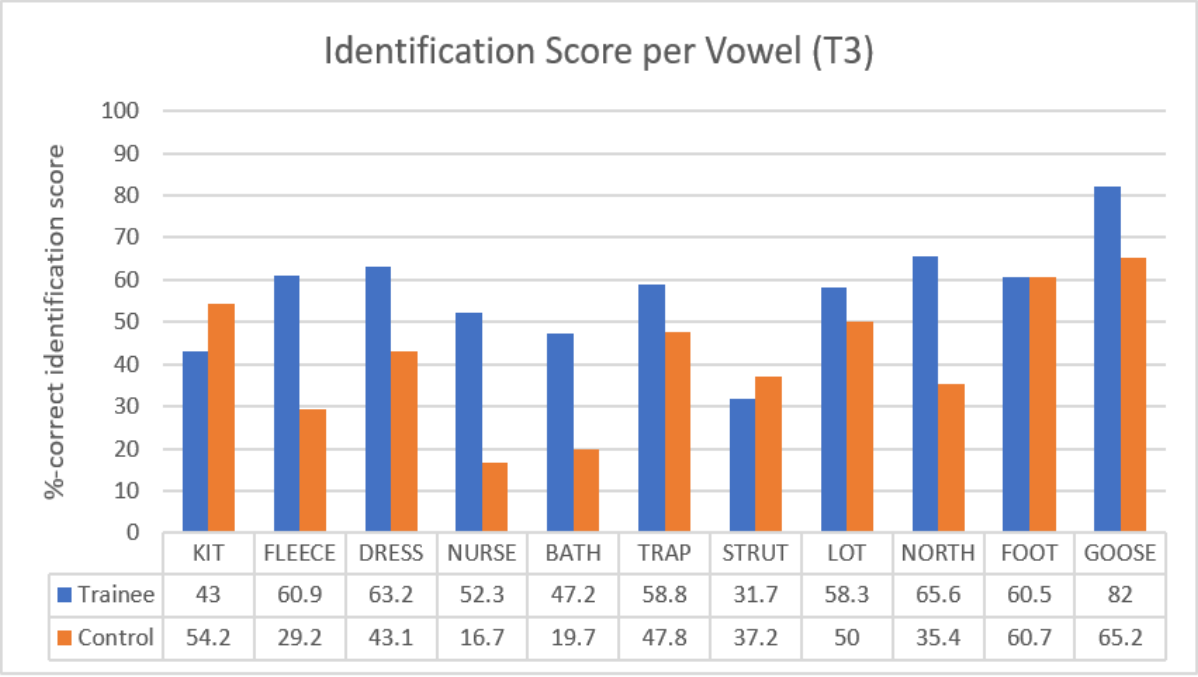
	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
Trainee	28.1	37	54.5	34.8	21.6	42.9	30.8	50	38.6	61.3	62.6
Control	34.4	29.2	45.5	15.5	15.2	43.3	34.6	43.3	31.3	53.3	54.5

Figure 8. Mean perceptual identification scores of Trainees and Controls in each vowel at T1



	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
Trainee	43.8	63	52.8	44.1	46.6	58.7	36.2	56.7	52.3	67.1	70.5
Control	49	28.1	43.9	22.6	24.2	45.6	30.8	56.7	42.7	60	65.2

Figure 9. Mean perceptual identification scores of Trainees and Controls in each vowel at T2



	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
Trainee	43	60.9	63.2	52.3	47.2	58.8	31.7	58.3	65.6	60.5	82
Control	54.2	29.2	43.1	16.7	19.7	47.8	37.2	50	35.4	60.7	65.2

Figure 10. Mean perceptual identification scores of Trainees and Controls in each vowel at T3

In order to assess whether these differences were significant, binomial logistic regressions using the formula *glmer (Response~Time+(1|Subject), data=data.frame, family= "binomial")* were conducted for each CYG group. These showed that while controls only improved in KIT across time (est.=0.629, p=0.376 at T2; est.=0.846, p=0.005 at T3), trainees improved significantly in most vowels at T2 and/or T3: KIT (est.=0.746, p=0.007 at T2;

est.=0.711, p=0.01 at T3), FLEECE (est.=1.428, p<0.001 at T2; est.=1.31, p<0.001 at T3), NURSE (est.=0.81, p=0.006 at T3), BATH (est.=1.33, p<0.001 at T2; est.=1.367, p<0.001 at T3), TRAP (est.=0.681, p=0.011 at T2; est.=0.677, p=0.012 at T3), NORTH (est.=0.682, p=0.015 at T2; est.=1.374, p<0.001 at T3), and GOOSE (est.=1.071, p=0.00334 at T3).

Furthermore, even though the two groups had no significant differences between them in the identification of each target vowel at T1 (formula: *glmer (Response~Group+(1|Subject), data=data.frame, family= "binomial")*), trainees identified some vowels better than controls at T2 (FLEECE: est.=2.093, p=0.036; NURSE: est.=1.097, p=0.028) and T3 (FLEECE: est.=1.83, p=0.02; NURSE: est.=1.905, p<0.001; BATH: est.=1.606, p=0.035; NORTH: est.=1.502, p=0.019; GOOSE: est.=0.906, p=0.027).

4.1.4 Generalisation Test

The post-test and retention test included a set of generalisation tests as well. Each generalisation test consisted of three new sets of stimuli: 98 stimuli with vowels in untrained contexts (/sVt/ and /dVt/) produced by known voices (New Contexts test), 36 stimuli with vowels produced by two new speakers in the known /bVt/ and /gVt/ contexts (New Speakers test), and 25 stimuli with vowels produced by the two new speakers in the untrained contexts (New Speakers and Contexts test). The %-correct identification scores of CYG trainees and controls in each of these tests are shown in **Figure 11**.

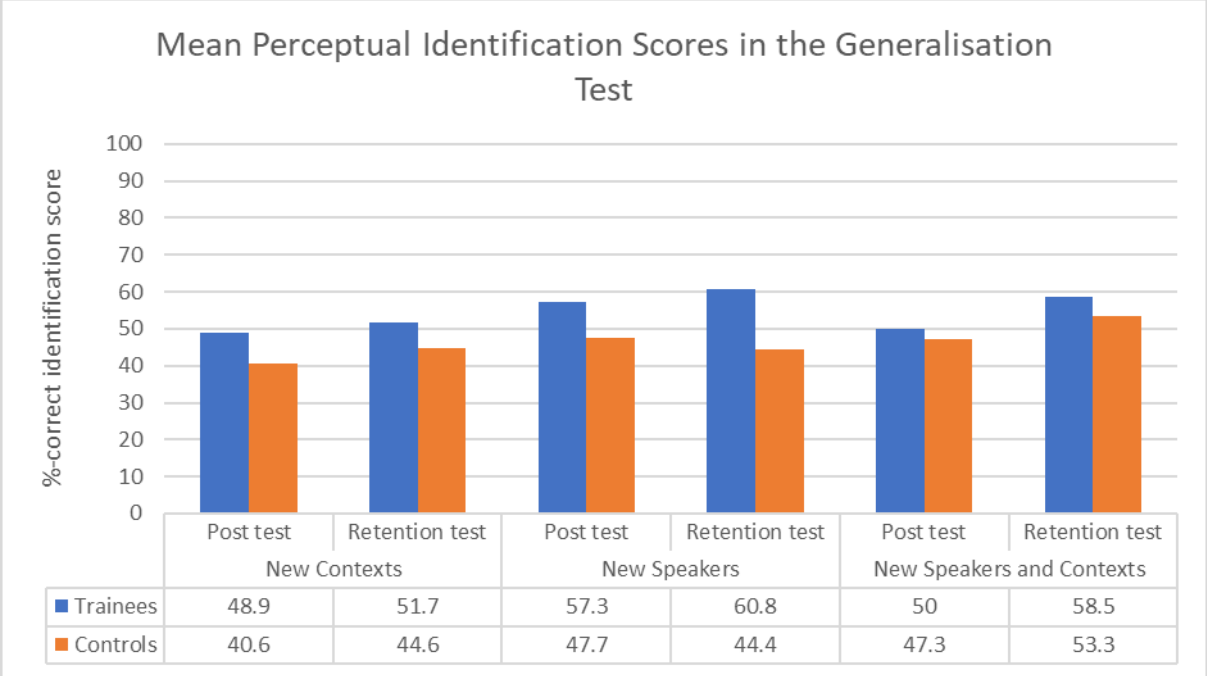


Figure 11. Mean perceptual identification scores of trainees and controls in each generalisation test at T2 and T3

The differences between groups at T2 and T3 as well as within-group differences across time were analysed using mixed-effects binary logistic regression. In the New Contexts and in the New Speakers and Contexts tests, the analysis showed that there were no significant differences between the two groups at either T2 or T3, and no significant differences between T2 and T3 identification for either group. In the New Speakers test, the analysis showed significant differences between the two groups at T3 only (est.=0.681, p=0.014), but differences in perceptual identification between T2 and T3 did not reach significance for either group.

Mixed-effects binary logistic regression was also used to compare each CYG group's performance in the post-test and retention test to their performance in each of the generalisation tests. The difference between trainees' performance at the post-test and retention test compared to their performance in the New Contexts test at T2 and T3, respectively, reached significance (est.=0.206, p=0.036 and est.=0.203, p=0.034, respectively), suggesting that they performed better in known than in new contexts. The lack of any other statistically significant differences at T2 indicates that participants' performance was largely unaffected by the inclusion of new speakers. Finally, there was a significant difference between controls' performance in the retention test and their performance in the New Speakers and Contexts test (est.=-0.495, p=0.006), where they performed better than in any other test and at any other time.

4.2 Acoustic Analysis – Duration

4.2.1 Pre-test (T1)

Table 6 presents the average duration (in ms) of each vowel as produced by learners and NE speakers at T1 in each task. As mentioned in 3.6, NE speakers only completed a wordlist-reading task. For the purposes of these analyses, the mean duration of each target vowel includes productions in both the /bVt/ and /gVt/ contexts, given the absence of any significant differences between them in either group, as determined through linear mixed-effects analyses with *Context* as the fixed and *Participant* and *Vowel* as the random effects. As can be seen in **Table 6**, vowels in the wordlist-reading task were consistently longer than vowels produced in the elicitation task. This was further confirmed through a linear mixed-effects analysis for each target vowel, with *Task* (Elicitation, Wordlist-reading) and *Subgroup* (CYG Trainee, CYG Control) as the fixed effects and *Subject* as the random effect (formula: *lmer*

*(Duration~Task*Subgroup+(1|Subject), data=data.frame*). The effect of *Task* reached significance in most cases: KIT (est.=31.857, p=0.001), FLEECE (est.=44.365, p<0.001), DRESS (est.=35.237, p=0.004), TRAP (est.=31.48, p=0.009), STRUT (est.=53.87, p<0.001), NORTH (est.=39.584, p<0.001), FOOT (est.=35.604, p=0.028). The effect of *Subgroup* did not reach significance in any of the comparisons. The interaction effect did not reach significance in any of the vowels either: KIT (est.=0.718, p=0.953), FLEECE (est.=-24.99, p=0.122), DRESS (est.=1.624, p=0.912), NURSE (est.=2.28, p=0.911), BATH (est.=6.936, p=0.719), TRAP (est.=-6.21, p=0.678), STRUT (est.=-22.15, p=0.146), LOT (est.=13.375, p=0.329), NORTH (est.=-14.115, p=0.272), FOOT (est.=4.015, p=0.843), GOOSE (est.=7.263, p=0.67).

VOWEL	**TRAINEES**		**CONTROLS**		**COMBINED CYG**		**NE SPEAKERS**
	Elicitation	*Wordlist*	*Elicitation*	*Wordlist*	*Elicitation*	*Wordlist*	W*ordlist*
KIT	108	140	98	130	103	136	128
FLEECE	126	146	110	153	119	149	153
DRESS	133	170	125	160	130	166	120
NURSE	162	174	163	173	162	173	212
BATH	182	190	182	182	182	186	231
TRAP	163	188	145	176	155	183	144
STRUT	146	178	137	186	142	182	125
LOT	143	174	144	162	143	169	123
NORTH	162	188	142	181	153	185	199
FOOT	135	173	119	156	128	166	129
GOOSE	138	170	131	156	135	163	177

Table 6. Mean duration (ms) of each target vowel produced in the bVt and gVt contexts by CYG Trainees, CYG Controls, both CYG groups combined and NE speakers at T1 in each task

In order to determine whether CYG learners produced any length differences in the members of a contrastive pair, the effects of *Vowel* and *Subgroup* were also examined (*formula: lmer (Duration~Vowel*Subgroup+(1|Subject), data=data.frame)*). The effect of *Subgroup* did not reach significance in any of the comparisons. The effect of *Vowel* was significant in DRESS-NURSE (est.=25.392, p=0.049), BATH-TRAP (est.=-24.723, p=0.026), BATH-STRUT (est.=-43.266, p<0.001) and TRAP-STRUT (est.=17.63, p=0.048) in the elicitation task, and in KIT-FLEECE (est.-21.492, p=0.003) and LOT-NORTH (est.=19.373, p=0.003) in the wordlist-reading task, suggesting that overall, CYG participants produced each of the vowels in these vowel pairs with different durations in the respective tasks, while consistently maintaining a shorter duration for the short vowel in each pair, compared to longer durations for the long vowels in each pair. A significant effect of the *Vowel*Subgroup* interaction was

observed in the LOT-NORTH pair in the elicitation task only (est.=22.815, p=0.023), where trainees produced LOT with a shorter duration than NORTH at T1, whereas controls had a slightly longer duration in LOT. No other interaction effects reached significance: KIT-FLEECE (est.=-7.783, p=0.543), DRESS-NURSE (est.=4.878, p=0.74), BATH-TRAP (est.=2.759, p=0.842), BATH-STRUT (est.=6.902, p=0.626), TRAP-STRUT (est.=-2.791, p=0.801), FOOT-GOOSE (est.=-7.651, p=0.651) in the elicitation task; KIT-FLEECE (est.=14.899, p=0.114), DRESS-NURSE (est.=-12.768, p=0.252), BATH-TRAP (est.=-0.18, p=0.989), BATH-STRUT (est.=-12.76, p=0.282), TRAP-STRUT (est.=15.708, p=0.209), LOT-NORTH (est.=-5.679, p=0.48277), FOOT-GOOSE (est.=-7.015, p=0.544) in the wordlist-reading task.

To assess whether the vowels in a vowel pair were produced with different lengths in each of the tasks, the effects of *Task* (Elicitation, Wordlist) and *Vowel* on duration in CYG speakers' productions at T1 were also examined using the formula *lmer (Duration~Task*Vowel+(1|Subject), data=data.frame)*. A significant effect of *Task* was found in most comparisons: KIT-FLEECE (est.=30.832, p<0.001), DRESS-NURSE (est.=35.996, p<0.001), TRAP-STRUT (est.=41.508, p<0.001), LOT-NORTH (est.=25.429, p<0.001) and FOOT-GOOSE (est.=38.825, p<0.001), indicating that vowels in the wordlist-reading task were produced with longer durations than in the elicitation task in most cases. A significant effect of *Vowel* was observed in KIT-FLEECE (est.=-15.853, p=0.022), DRESS-NURSE (est.=32.611, p<0.001), BATH-TRAP (est.=-26.667, p=0.002) and BATH-STRUT (est.=-40.383, p<0.001), demonstrating that irrespective of the task, one of the two vowels (the long vowel in these cases) in these pairs was produced as consistently longer than the other.

Finally, the interaction effect was only significant in the DRESS-NURSE (est.=-25.001, p=0.032) and BATH-STRUT (est.=36.877, p=0.001) pairs, indicating that the effect of *Task* was different for each of the two vowels in these pairs. More specifically, the short vowels in these pairs, i.e. DRESS and STRUT, were considerably longer in the wordlist-reading task (166ms and 182ms, respectively when averaged across all CYG participants) compared to the elicitation task (130ms and 142ms, respectively); the long vowels NURSE and BATH were also longer in the wordlist-reading task (173ms and 186ms, respectively) compared to the elicitation task (162ms and 182ms), but the difference was much smaller. No other interaction effects reached significance in this comparison: KIT-FLEECE (est.=1.699, p=0.862), BATH-TRAP (est.=21.772,

p=0.05), TRAP-STRUT (est.=-13.871, p=0.166), LOT-NORTH (est.=6.317, p=0.481), FOOT-GOOSE (est.=-9.363, p=0.442).

4.2.2 Post-test (T2)

Table 7 presents the average duration of each vowel produced by learners in each task at T2. Similar analyses were conducted to assess vowel duration at T2. As with T1 data, vowels in the wordlist-reading task were consistently longer than vowels produced in the elicitation task at T2. In the examination of the effects of *Task* (Elicitation, Wordlist) and *Subgroup* (CYG Trainee, CYG Control), the effect of *Task* reached significance in most cases: KIT (est.=24.53, p=0.027), DRESS (est.=30.808, p=0.004), TRAP (est.=31.729, p=0.011), LOT (est.=25.499, p=0.016), NORTH (est.=35.73, p=0.016), FOOT (est.=30.771, p=0.046) and GOOSE (est.=31.37, p=0.015). The effect of *Subgroup* and of the interaction *Task*Subgroup* did not reach significance in any of the comparisons, similarly to T1. Interaction effects were found to be as follows: KIT (est.=-20.942, p=0.149), FLEECE (est.=0.453, p=0.981), DRESS (est.=-6.631, p=0.621), NURSE (est.=-8.464, p=0.611), BATH (est.=2.195, p=0.877), TRAP (est.=-8.594, p=0.583), STRUT (est.=13.46, p=0.422), LOT (est.=-9.381, p=0.491), NORTH (est.=-23.76, p=0.195), FOOT (est.=-0.914, p=0.962), GOOSE (est.=-11.689, p=0.47).

VOWEL	TRAINEES		CONTROLS	
	Elicitation	*Wordlist*	*Elicitation*	*Wordlist*
KIT	111	116	114	138
FLEECE	119	146	118	144
DRESS	123	147	134	165
NURSE	165	167	166	180
BATH	172	192	171	193
TRAP	150	173	149	179
STRUT	136	167	162	177
LOT	142	158	144	168
NORTH	166	178	138	172
FOOT	126	155	118	150
GOOSE	130	150	123	153

Table 7. Mean duration of each target vowel produced in the /bVt/ and /gVt/ contexts by CYG trainees and CYG controls at T2

The effects of *Vowel* and *Subgroup* (CYG Trainee, CYG Control) were also examined in each of the two tasks, to determine whether CYG learners produced any length differences in the members of a contrastive pair. The effect of *Subgroup* did not reach significance in any of the comparisons, similarly to T1 data. In the elicitation task, there was a significant effect of

Vowel in DRESS-NURSE (est.=24.87, p=0.038), where the long vowel was significantly longer than the short vowel in the pair, and of the *Vowel*Subgroup* interaction in BATH-STRUT (est.=-31.541, p=0.025), where the long vowel was almost 36ms longer than the short vowel for trainees, but only 8.9ms longer for the controls. In the wordlist-reading task, none of the comparisons yielded any significant effects of the IVs. No other significant interaction effects were observed in either task: KIT-FLEECE (est.=1.448, p=0.891), DRESS-NURSE (est.=17.14, p=0.222), BATH-TRAP (est.=-10.34, p=0.466), TRAP-STRUT (est.=19.674, p=0.081), LOT-NORTH (est.=25.209, p=0.062), FOOT-GOOSE (est.=-4.606, p=0.75) in the elicitation task; KIT-FLEECE (est.=-18.911, p=0.187), DRESS-NURSE (est.=7.68, p=0.503), BATH-TRAP (est.=-6.106, p=0.61), BATH-STRUT (est.=-11.176, p=0.391), TRAP-STRUT (est.=6.525, p=0.628), LOT-NORTH (est.=14.2, p=0.223), FOOT-GOOSE (est.=-8.182, p=0.49) in the wordlist-reading task.

Finally, the analysis of the effects of *Task* and *Vowel* showed a significant effect of *Task* for both groups in KIT-FLEECE (est.=26.651, p=0.031 for trainees; est.=26.534, p=0.023 for controls), DRESS-NURSE (est.=24.177, p=0.007 for trainees; est.=29.79, p=0.007 for controls) and FOOT-GOOSE (est.=30.276, p=0.008 for trainees; est.=31.247, p=0.027 for controls), indicating shorter productions for all these vowels in the elicitation than in the wordlist-reading task. An effect of *Task* was also observed in TRAP-STRUT for trainees (est.=30.259, p=0.005) and in LOT-NORTH (est.=25.146, p=0.01) for controls. Finally, the effect of *Vowel* reached significance in DRESS-NURSE for both groups (est.=43.534, p<0.001 for trainees; est.=33.09, p=0.019 for controls) and in BATH-TRAP (est.=-23.123, p=0.021), BATH-STRUT (est.=-36.827, p<0.001) and LOT-NORTH (est.=25.112, p=0.028) for trainees only, demonstrating that irrespective of task, the short vowels in these pairs were produced with significantly shorter durations than the long vowels.

4.2.3 Comparison of T1 and T2 Durations

To assess whether there were any differences between the two CYG groups at T1 and T2, the effects of *Time* (T1, T2) and *Subgroup* (CYG Trainee, CYG Control) on the duration of individual vowels in each of the two tasks were examined in a linear mixed-effects model analysis (formula: *lmer (Duration~Time*Subgroup+(1|Subject), data=data.frame)*). In the elicitation task, there was a significant effect of *Time* and the *Time*Subgroup* interaction only in STRUT (est.=23.303, p=0.013 and est.=-34.689, p=0.004, respectively). More specifically, the trainees shortened the vowel by almost 10ms at T2, whereas the controls lengthened the vowel by

more than 25ms at T2. In the wordlist-reading task, there was a significant effect of the *Time*Subgroup* interaction in KIT (est.=-28.011, p=0.018), DRESS (est.=-27.777, p=0.008) and LOT (est.=-20.382, p=0.032). The same pattern was observed here: the trainees shortened these vowels at T2, whereas the controls lengthened the vowels, albeit not significantly. No other interaction effects reached significance in either task: KIT (est.=-14.206, p=0.194), FLEECE (est.=-20.2, p=0.104), DRESS (est.=-19.069, p=0.07), NURSE (est.=-5.643, p=0.749), BATH (est.=1.533, p=0.928), TRAP (est.=-13.854, p=0.267), LOT (est.=-1.119, p= 0.93), NORTH (est.=3.184, p=0.822), FOOT (est.=-7.196, p=0.652), GOOSE (est.=-0.444, p=0.979) in the elicitation task; FLEECE (est.=3.045, p=0.791), NURSE (est.=-18.955, p=0.081), BATH (est.=-10.169, p=0.361), TRAP (est.=-15.212, p=0.251), STRUT (est.=-3.684, p=0.795), NORTH (est.=0.619, p=0.945), FOOT (est.=-11.963, p=0.486), GOOSE (est.=-17.089, p=0.075) in the wordlist-reading task.

To assess differences between contrastive vowels at T1 and T2 in each of the two CYG groups, the effects of *Time* (T1, T2) and *Vowel* were examined using the formula *lmer (Duration~Time*Vowel+(1|Subject), data=data.frame)*. For trainees in the elicitation task, only the effect of *Vowel* reached significance, and only in the vowel pairs DRESS-NURSE (est.=30.533, p<0.001), BATH-TRAP (est.=-21.514, p=0.027), BATH-STRUT (est.=-35.433, p<0.001) and LOT-NORTH (est.=19.357, p=0.048); in these cases, the long vowel in the pair was significantly longer than the short vowel at T1, and the difference was maintained or increased at T2. The effect of *Time* and *Time*Vowel* did not reach significance in any of the comparisons for this group in the elicitation task. Interaction effects were found to be as follows: KIT-FLEECE (est.=8.533, p=0.526), DRESS-NURSE (est.=11.727, p=0.318), BATH-TRAP (est.=-1.502, p=0.907), BATH-STRUT (est.=-0.44, p=0.973), TRAP-STRUT (est.=-1.757, p=0.878), LOT-NORTH (est.=6.244, p=0.646), FOOT-GOOSE (est.=0.422, p=0.975).

As concerns the wordlist-reading task, only a significant effect of *Time* and only for the vowels DRESS-NURSE (est.=-22.878, p=0.003) and LOT-NORTH (est.=-15.915, p=0.042) was found for trainees. The initial duration of the four vowels was similar, with the exception of NORTH (170ms, 174ms, 174ms and 188ms, respectively), and all four vowels became shorter at T2, albeit in differing degrees (23ms in DRESS, 7ms in NURSE, 16ms in LOT, 10ms in NORTH), resulting in a slightly increased distance between the short and long vowel in each pair. No significant effect of the interaction was observed in this task: KIT-FLEECE (est.=-20.715,

p=0.115), DRESS-NURSE (est.=18.001, p=0.115), BATH-TRAP (est.=-15.842, p=0.198), BATH-STRUT (est.=-15.797, p=0.259), TRAP-STRUT (est.=-1.172, p=0.93), LOT-NORTH (est.=6.245, p=0.564), FOOT-GOOSE (est.=-2.024, p=0.874).

In the control group, there was a significant effect of *Vowel* in BATH-TRAP (est.=-24.615, p=0.030), BATH-STRUT (est.=-42.499, p<0.001) and TRAP-STRUT (est.=17.276, p=0.047) in the elicitation task, where BATH was shortened at T2, and TRAP and STRUT became longer in these cases, moving further away from NE durations. In the same task, a significant effect of *Time* was also found in TRAP-STRUT (est.=23.155, p=0.011), where STRUT was lengthened at T2, and an effect of the *Time*Vowel* interaction was observed in BATH-STRUT (est.=37.73, p=0.017), in which case BATH was shortened by more than 10ms, while STRUT was lengthened by more than 25ms at T2. No other interaction effects reached significance: KIT-FLEECE (est.=4.065, p=0.692), DRESS-NURSE (est.=1.617, p=0.913), BATH-TRAP (est.=13.555, p=0.366), TRAP-STRUT (est.=-22.721, p=0.059), LOT-NORTH (est.=6.05, p=0.533), FOOT-GOOSE (est.=-3.310, p=0.851).

In the wordlist-reading task, only a significant effect of *Vowel* was observed and only in FLEECE-KIT (est.=-21.745, p=0.003) and LOT-NORTH (est.=20.203, p<0.001). As opposed to trainees, however, controls decreased the distance in duration between the two vowels of each of these pairs, by lengthening the short vowels and shortening the long vowels in the pair at T2, thereby moving away from NE durations. The interaction effect did not reach significance for controls either in this task: KIT-FLEECE (est.=13.13, p=0.182), DRESS-NURSE (est.=3.941, p=0.718), BATH-TRAP (est.=-11.016, p=0.404), BATH-STRUT (est.=-20.587, p=0.054), TRAP-STRUT (est.=9.062, p=0.429), LOT-NORTH (est.=-15.078, p=0.062), FOOT-GOOSE (est.=3.054, p=0.793).

4.2.4 Comparison with NE Speakers and CYG Vowels

The examination of vowel duration in the productions of NE speakers showed that even though this is a secondary cue in vowel production and perception in English, NE speakers maintained duration differences between the members of all contrastive pairs, as indicated by the effect of *Vowel* on duration in each contrast: KIT-FLEECE (est.=-24.763, p=0.001), DRESS-NURSE (est.=93.81, p<0.001), BATH-TRAP (est.=-86.652, p<0.001), BATH-STRUT (est.=-100.537, p<0.001), TRAP-STRUT (est.=16.588, p=0.008), LOT-NORTH (est.=75.071, p<0.001), FOOT-GOOSE (est.=48.1, p<0.001).

To assess whether vowel duration was different between NE speakers and CYG learners in each target vowel in the wordlist-reading task at T1, the effect of *Group* (NE, CYG both groups) was examined in a linear mixed-effects analysis for each target vowel, with *Duration* as the DV, *Group* as the fixed effect and *Subject* as the random effect. A significant effect of *Group* was found in most comparisons: DRESS (est.=-47.5, p=0.001), NURSE (est.=37.518, p=0.002), BATH (est.=45.523, p<0.001), TRAP (est.=-40.035, p=0.002), STRUT (est.=-59.728, p<0.001), LOT (est.=-46.373, p<0.001) and FOOT (est.=-41.07, p=0.014). This means that the two groups produced all vowels with the exception of KIT, FLEECE, NORTH and GOOSE with different durations. More specifically, CYG learners produced the vowels in KIT-FLEECE and LOT-NORTH with distinct durations (effect of *Vowel* in 4.2.1), which were close to NE values with the exception of LOT. GOOSE also had a duration that approximated NE values; however, CYG learners produced the vowels in each of the vowel sets DRESS-NURSE, BATH-TRAP-STRUT and FOOT-GOOSE with similar durations, which were intermediate between the respective long and short vowels in NE speech, i.e. they had longer durations in short vowels and shorter durations in long vowels compared to NE speakers.

The same analysis was conducted for the elicitation task at T1, even though the productions of NE speakers were collected in a wordlist-reading task. In this case, there was a significant effect of *Group* in KIT (est.=24.116, p<0.001), FLEECE (est.=32.762, p<0.001), NURSE (est.=47.413, p<0.001), BATH (est.=49.625, p<0.001), NORTH (est.=44.255, p<0.001) and GOOSE (est.=42.678, p=0.002), showing that the generally shorter durations of vowels in the elicitation task may have helped learners approximate NE durations in all short vowels except for KIT. However, the differences in long vowels may also be attributed to a task effect that cannot be further examined.

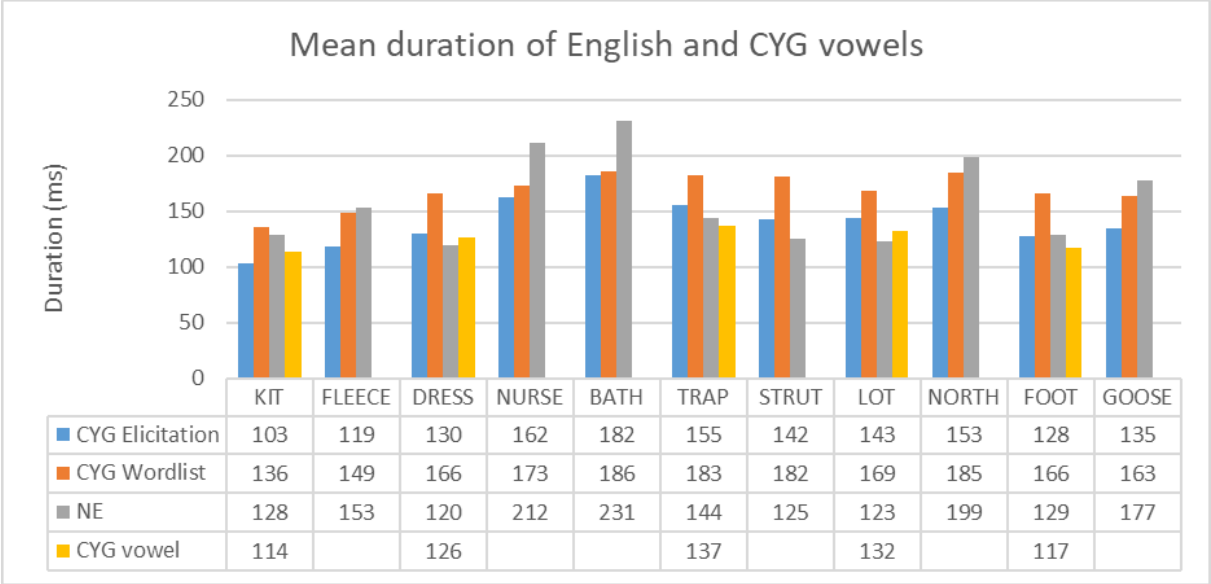
To assess whether there were significant differences in the duration of individual vowels between NE speakers and trainees in the wordlist-reading task at T2, the effect of *Group* (NE, CYG Trainees) was examined using the same formula for each target vowel. This yielded a significant effect of *Group* in NURSE (est.=43.657, p=0.003), BATH (est.=39.861, p=0.002), STRUT (est.=-41.41, p=0.02) and LOT (est.=-35.28, p=0.019). In these cases, the long vowels of NE speakers were significantly longer, and the short vowels significantly shorter than trainees' productions. In all other cases (KIT, FLEECE, DRESS, TRAP, NORTH, FOOT and GOOSE), the differences between NE speakers and CYG trainees did not reach significance. For controls,

there was a significant effect of *Group* in most vowels in the wordlist-reading task: DRESS (est.=-46.151, p=0.002), NURSE (est.=32.65, p=0.041), BATH (est.=38.84, p=0.028), TRAP (est.=-35.671, p=0.012), STRUT (est.=-54.48, p=0.005), LOT (est.=-43.47, p=0.011) and NORTH (est.=25.683, p=0.049). This means that CYG controls approximated NE durations in the high vowels KIT, FLEECE, FOOT and GOOSE only.

In the elicitation task at T2, the effect of *Group* on duration reached significance in the same, long vowels for trainees and controls: NURSE (for trainees: est.=46.65, p=0.007; for controls: est.=44.91, p=0.021), BATH (for trainees: est.=59.11, p<0.001; for controls: est.=60.79, p=0.006), NORTH (for trainees: est.=32.658, p=0.008; for controls: est.=61.758, p<0.001) and GOOSE (for trainees: est.=46.99, p=0.009; for controls: est.=52.05, p=0.021).

Finally, the duration of CYG vowels as produced by the learners in a wordlist-reading task was analysed. The mean duration of CYG vowels embedded in the ˈbVtV context were as follows: /i/=114ms, /e/=126ms, /a/=137ms, /o/=132ms and /u/=117ms. In order to determine whether duration varied as an effect of *Vowel* in CYG vowels, an one-way ANOVA was conducted, since there were no missing values in this set (formula: *aov(Duration~Vowel, data=data.frame)*. No significant effect of *Vowel* was found (F=2.356, p=0.063), suggesting that CYG vowels were not differentiated by duration.

Figure 12 shows the mean duration of each vowel produced by CYG learners in each task in the /bVt/ and /gVt/ contexts at T1, the NE durations, as well as the L1 vowel productions of CYG speakers, shown next to the closest L2 vowel in terms of duration. Evidently, CYG vowels are shorter than L2 long vowels and approximate the durations of the corresponding L2 short vowels in each set. This was confirmed through linear mixed-effects model analysis where each CYG vowel was compared with the corresponding L2 vowel produced by the NE speakers in separate models (FLEECE-i, KIT-i, DRESS-e, NURSE-e, BATH-a, TRAP-a, STRUT-a, LOT-o, NORTH-o, FOOT-u, GOOSE-u) with *Duration* as the DV, *Vowel* as the fixed effect and *Subject* as the random effect. A significant effect of *Vowel* was observed in the comparisons between each CYG vowel and the corresponding long L2 vowel (FLEECE-i: est.=-39.566; NURSE-e: est.=85.087; BATH-a: est.=94.421; NORTH-o: est.=-65.991; GOOSE-u: est.=-60.673; p<0.001 in all comparisons), whereas no effect was found in the comparisons of CYG vowels with the corresponding short vowels KIT, DRESS, TRAP, STRUT, LOT and FOOT.



	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
CYG Elicitation	103	119	130	162	182	155	142	143	153	128	135
CYG Wordlist	136	149	166	173	186	183	182	169	185	166	163
NE	128	153	120	212	231	144	125	123	199	129	177
CYG vowel	114		126			137		132		117	

Figure 12. Mean duration of target vowels by CYG learners (/bVt/-/gVt/ contexts) in each task at T1 and by NE speakers, and mean duration of corresponding CYG vowels

4.3 Acoustic Analysis – Spectral Characteristics

4.3.1 Pre-test (T1)

Before proceeding with any comparisons, the Pillai score was calculated to assess the overlap in vowels produced in the /bVt/ and /gVt/ contexts (formula: *manova (cbind(F1, F2)~Context, data=data.frame)*). **Table 8** shows the mean formant values (Hz) of each vowel in each context as produced by all CYG learners in both the wordlist-reading and elicitation tasks, and the Pillai scores obtained for each comparison between them. **Figure 13** shows the plotted vowels for each of the contexts, demonstrating the amount of overlap between them. As can be seen, there is a high degree of overlap in the vowels produced in each context, with the highest Pillai score being 0.2342 for DRESS. Therefore, analyses were conducted with combined /bVt/ and /gVt/ mean values.

VOWEL	CONTEXT				PILLAI SCORE	P-VALUE
	/bVt/		/gVt/			
KIT	306	1769	305	1788	0.0044	
FLEECE	301	1746	289	1822	0.1588	0.017
DRESS	422	1545	402	1595	0.2342	<0.001
NURSE	419	1391	411	1488	0.2072	0.019
BATH	503	1175	520	1227	0.0569	
TRAP	551	1261	544	1315	0.116	0.046
STRUT	533	1234	514	1289	0.1657	0.013
LOT	416	1039	421	1087	0.1476	0.018
NORTH	400	1027	403	1047	0.0223	
FOOT	326	1077	321	1144	0.1165	
GOOSE	321	1077	327	1017	0.0961	

Table 8. Mean, Lobanov-normalised F1 and F2 values (Hz) of each vowel in each context, overlap between them (Pillai score) and significance of result

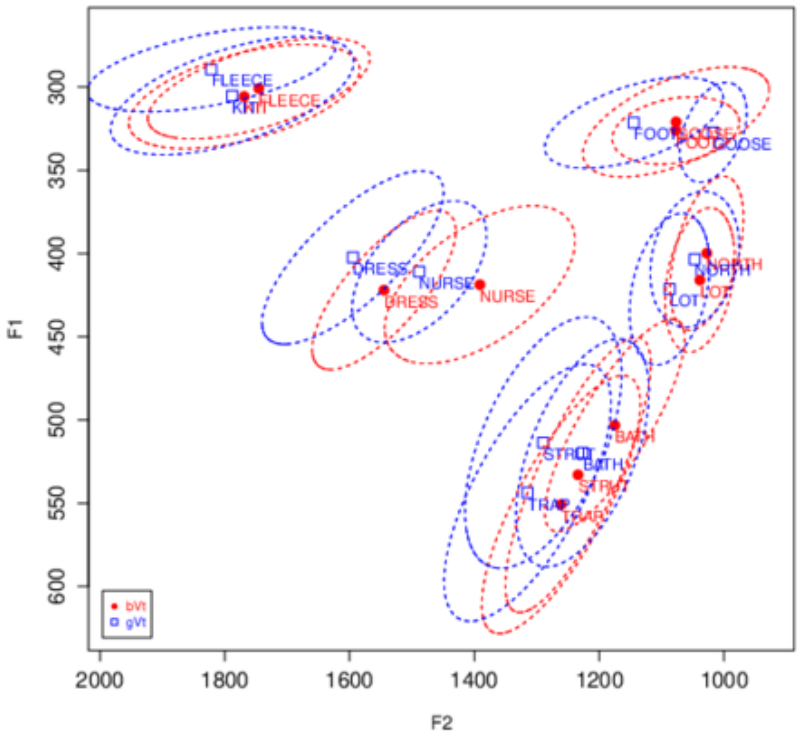


Figure 13. F1xF2 plot of target vowels in /bVt/ and /gVt/ contexts

Table 9 shows the mean F1 and F2 values of vowels produced by trainees and controls in each of the tasks. A within-group comparison was conducted to assess the overlap of vowels in the elicitation compared to the wordlist-reading task productions of participants in each CYG group (formula: *manova (cbind(F1, F2)~Task, data=data.frame)*). As shown in **Table 10**, productions between the two tasks were very similar in the trainee group, with the least overlap in TRAP and GOOSE; the same applied to controls, with the exception of NURSE, which

had a high Pillai score. **Figure 14** shows the F1xF2 plots of the productions of each group in each task.

VOWEL	TRAINEES				CONTROLS			
	Elicitation		*Wordlist*		*Elicitation*		*Wordlist*	
	F1	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
KIT	322	1785	308	1848	291	1717	294	1739
FLEECE	289	1782	307	1872	291	1735	291	1733
DRESS	420	1584	432	1609	386	1548	403	1522
NURSE	411	1417	432	1425	404	1367	402	1528
BATH	499	1219	529	1205	488	1151	518	1218
TRAP	548	1359	570	1276	531	1252	531	1237
STRUT	507	1293	553	1279	501	1237	524	1220
LOT	421	1088	438	1064	407	1059	403	1031
NORTH	400	1052	418	1019	390	1051	394	1028
FOOT	333	1130	326	1157	319	1089	309	1067
GOOSE	335	1156	323	1006	324	1045	311	1008

Table 9. Mean Lobanov-normalised F1 and F2 values (in Hz) of each vowel produced by the two groups in each task at T1

VOWEL	TRAINEES		CONTROLS		BOTH GROUPS	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*
KIT	0.1131		0.01		0.059	
FLEECE	0.1313		0.0006		0.038	
DRESS	0.0131		0.1275		0.035	
NURSE	0.0565		0.3832		0.0485	
BATH	0.0958		0.0931		0.055	
TRAP	0.2873	0.009	0.0128		0.1534	0.016
STRUT	0.1416		0.0991		0.123	0.043
LOT	0.0748		0.0782		0.0687	
NORTH	0.1154		0.1046		0.103	
FOOT	0.0448		0.0536		0.0358	
GOOSE	0.2444	0.046	0.1224		0.1668	0.022

Table 10. Comparison between wordlist-reading and elicitation productions within each group and for both groups together

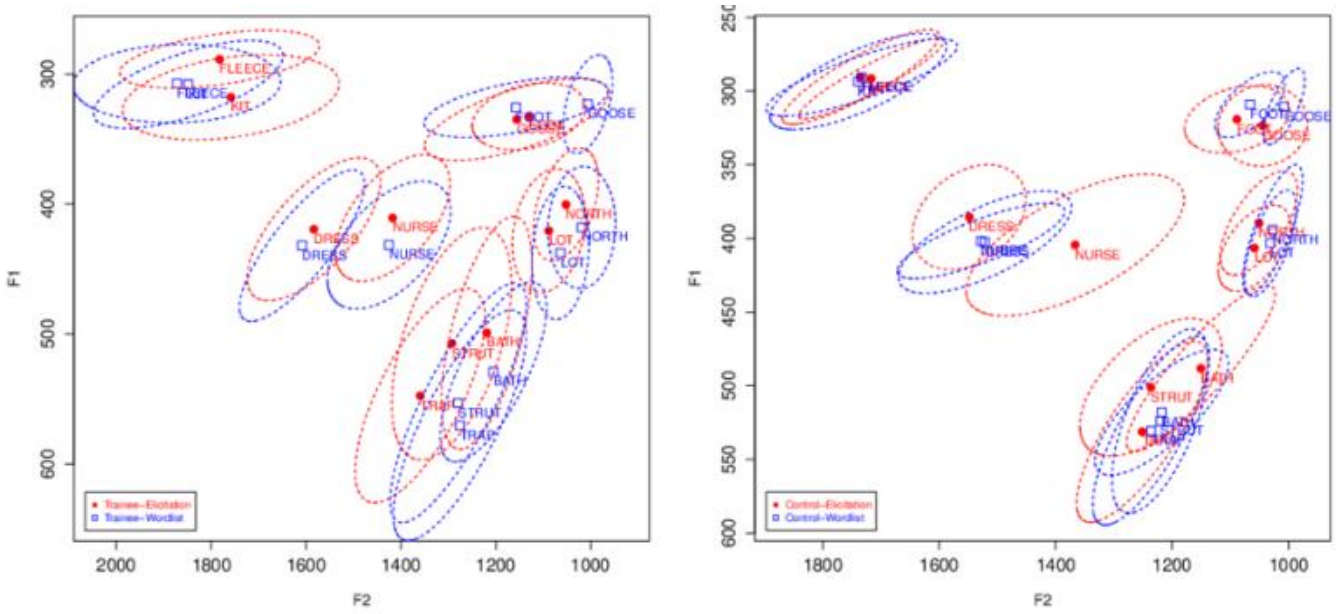


Figure 14. F1xF2 plot of target vowels produced by trainees (left) and controls (right) in each task

A comparison between the productions of trainees and controls in each of the tasks (formula: *manova (cbind(F1, F2)~Subgroup, data=data.frame)*) showed that participants in the two groups produced the vowels similarly in most cases, as suggested by the Pillai scores reported in **Table 11**. The least overlap in the elicitation task was found in KIT, followed by TRAP, while in the wordlist-reading task, there was a high Pillai score in NURSE, demonstrating that the two groups had differences in the production of this vowel. **Figure 15** shows the F1xF2 plots of target vowels produced by trainees compared to controls in each of the tasks.

VOWEL	**ELICITATION**		**WORDLIST**	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT	0.2197	0.045	0.0875	
FLEECE	0.0392		0.1463	
DRESS	0.1231		0.1016	
NURSE	0.0288		0.4058	0.012
BATH	0.1227		0.0431	
TRAP	0.2165		0.0644	
STRUT	0.0568		0.0574	
LOT	0.0564		0.1641	
NORTH	0.0246		0.0877	
FOOT	0.0864		0.1344	
GOOSE	0.1208		0.0581	

Table 11. Comparison between trainees and controls' production of vowels in each of the two tasks

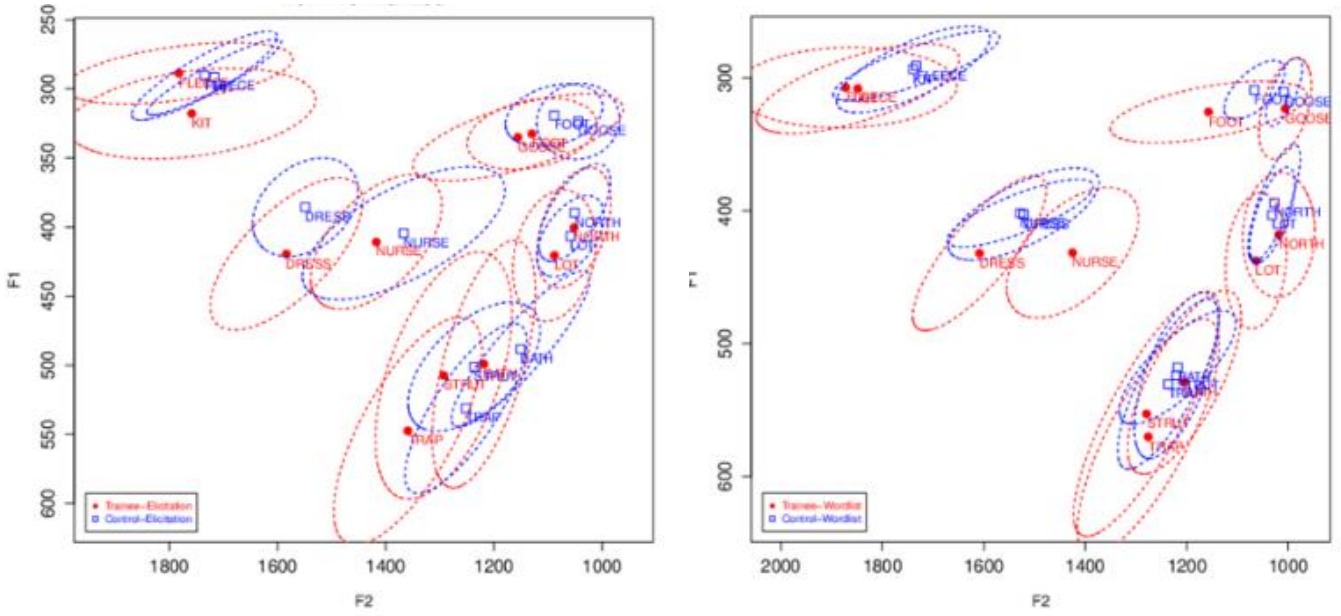


Figure 15. F1xF2 plot of target vowels produced by trainees and controls in the elicitation (left) and wordlist-reading (right) task

The examination of the vowels in a vowel pair at T1 in each task when both CYG groups' productions were combined (*manova (cbind(F1, F2)~Vowel, data=data.frame)*) showed considerable overlap between contrastive vowels (**Table 12**), with the highest Pillai score found in DRESS-NURSE in the elicitation task. This is also evident in the plot in **Figure 16**. Given some overlap observed between BATH and LOT, the Pillai score was also calculated for this pair. The overlap between contrastive vowels as produced by each CYG group in each task separately are shown in **Table 13**.

VOWEL PAIR	ELICITATION		WORDLIST	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT-FLEECE	0.1276	0.033	0.0031	
DRESS-NURSE	0.3994	<0.001	0.1827	0.011
BATH-TRAP	0.1973	0.011	0.0499	
BATH-STRUT	0.1225		0.0349	
TRAP-STRUT	0.0581		0.012	
LOT-NORTH	0.0621		0.0583	
FOOT-GOOSE	0.0063		0.238	
BATH-LOT	0.3677	<0.001	0.5389	<0.001

Table 12. Overlap between vowels of contrastive pairs as produced by both CYG groups at T1

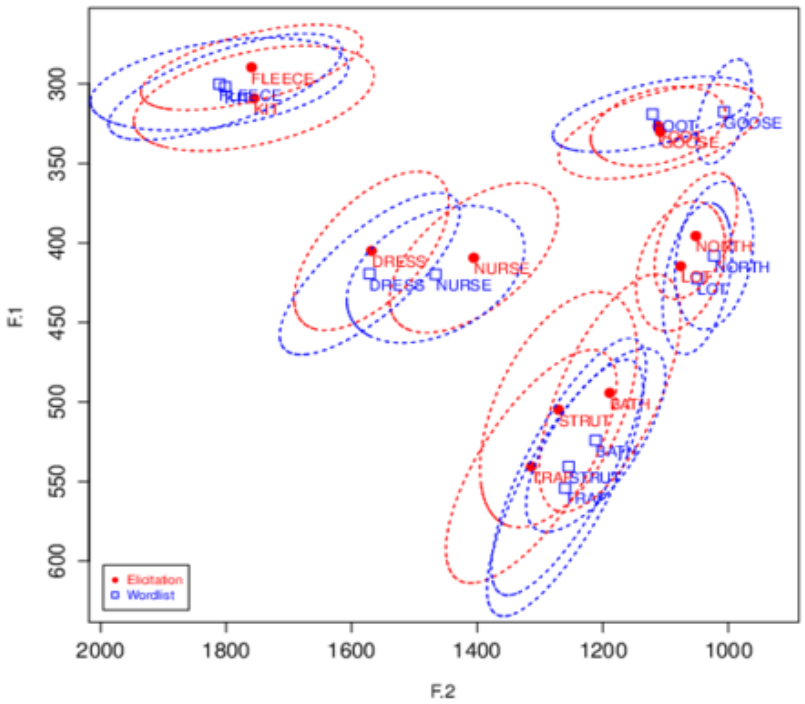


Figure 16. F1xF2 plot of the overlap of target vowels produced by CYG learners in the wordlist-reading and elicitation tasks at T1

VOWEL PAIR	TRAINEES				CONTROLS			
	Elicitation		*Wordlist*		*Elicitation*		*Wordlist*	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*
KIT-FLEECE	0.2994	0.01	0.0048		0.0372		0.0027	
DRESS-NURSE	0.3711	0.002	0.5068	<0.001	0.4561	0.019	0.0022	
BATH-TRAP	0.2632	0.035	0.0804		0.1638		0.0145	
BATH-STRUT	0.1157		0.0955		0.1444		0.0047	
TRAP-STRUT	0.0730		0.0318		0.1028		0.0087	
LOT-NORTH	0.0838		0.1336		0.0812		0.0177	
FOOT-GOOSE	0.0071		0.26	0.036	0.0861		0.4215	0.013
BATH-LOT	0.4128	0.003	0.4556	<0.001	0.5494	0.002	0.6977	<0.001

Table 13. Overlap between vowels of contrastive pairs as produced by trainees and controls in each task at T1

4.3.2 Post-test (T2)

Table 14 presents the mean, Lobanov-normalised formant values of the productions of trainees and controls in the two tasks at T2.

VOWEL	TRAINEES				CONTROLS			
	Elicitation		*Wordlist*		*Elicitation*		*Wordlist*	
	F1	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
KIT	323	1823	309	1860	298	1729	291	1757
FLEECE	307	1849	299	1897	293	1686	282	1714
DRESS	410	1545	435	1602	394	1557	400	1563
NURSE	403	1403	422	1419	409	1473	396	1508
BATH	498	1206	536	1205	483	1177	526	1215
TRAP	541	1353	577	1317	515	1243	539	1270
STRUT	532	1300	567	1268	490	1220	518	1248
LOT	415	1087	432	1075	406	1049	391	1045
NORTH	412	1058	396	1021	407	1047	381	1016
FOOT	331	1191	331	1067	325	1082	319	1019
GOOSE	331	1062	329	993	327	1083	314	951

Table 14. Mean Lobanov-normalised F1 and F2 values (in Hz) of each vowel produced by the two groups in each task at T2

A within-group comparison was conducted to assess the overlap of vowels in the elicitation compared to the wordlist-reading productions of participants within each group (formula: *manova (cbind(F1, F2)~Task, data=data.frame)*). As shown in **Table 15**, productions between the two tasks were very similar for both groups; the only exception was GOOSE in the control group, where wordlist-reading and elicitation productions had a high Pillai score and therefore very small overlap. **Figure 17** shows the F1xF2 plots of the productions of each group in each task at T2.

VOWEL	TRAINEES		CONTROLS	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT	0.0577		0.0881	
FLEECE	0.0631		0.2394	
DRESS	0.0575		0.0114	
NURSE	0.027		0.0853	
BATH	0.1271		0.1362	
TRAP	0.1904	0.047	0.0321	
STRUT	0.1411		0.0343	
LOT	0.0467		0.0483	
NORTH	0.0817		0.1257	
FOOT	0.197		0.1005	
GOOSE	0.1599		0.6166	<0.001

Table 15. Comparison between wordlist-reading and elicitation productions within each group (T2)

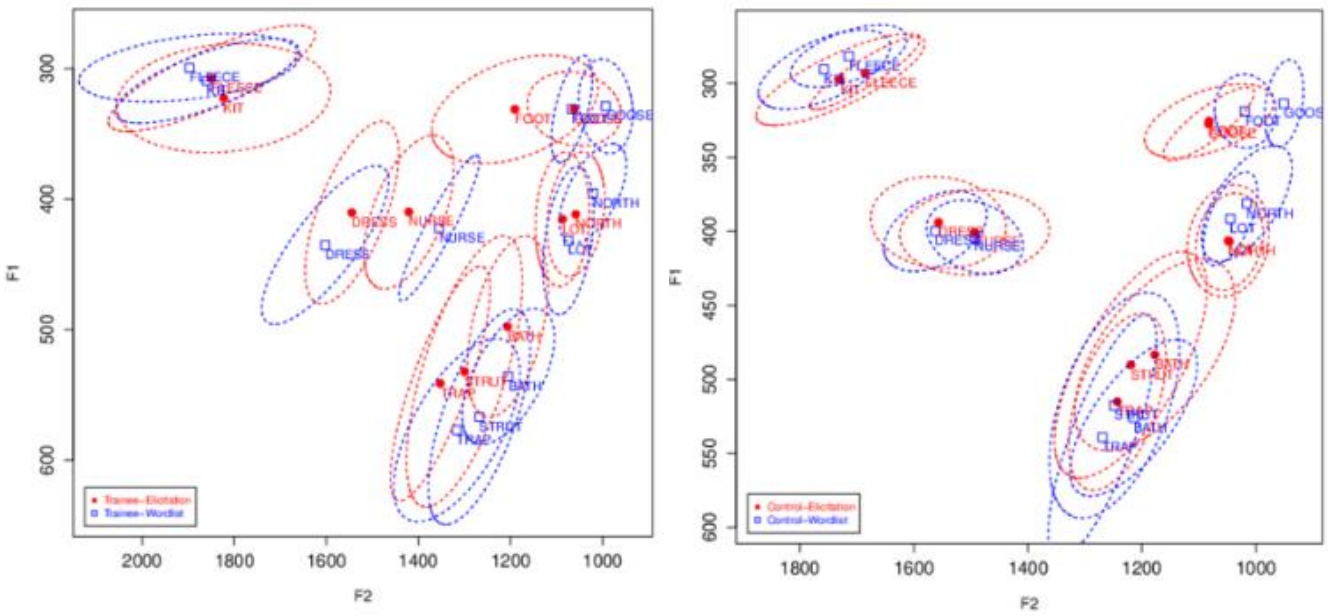


Figure 17. F1xF2 plot of target vowels produced by trainees (left) and controls (right) in each task at T2

A comparison between the productions of trainees and controls (formula: *manova (cbind(F1, F2)~Subgroup, data=data.frame)*) showed a large amount of overlap in most vowels in both tasks (**Table 16**). In the elicitation task, the least overlap was found in TRAP, followed by FLEECE, with most other vowels having a Pillai score close to 0. In the wordlist-reading task, the least overlap was found in NURSE, followed by FLEECE, GOOSE and FOOT. **Figure 18** shows the F1xF2 plots of target vowels produced by trainees compared to controls in each of the tasks at T2.

VOWEL	ELICITATION		WORDLIST	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT	0.1227		0.1058	
FLEECE	0.2136		0.2116	
DRESS	0.0369		0.1335	
NURSE	0.0731		0.3608	
BATH	0.02		0.0276	
TRAP	0.3021	0.011	0.0624	
STRUT	0.0908		0.102	
LOT	0.0908		0.1688	
NORTH	0.0102		0.0436	
FOOT	0.117		0.2032	
GOOSE	0.0204		0.2066	

Table 16. Comparison between trainees and controls' production of vowels in each of the two tasks (T2)

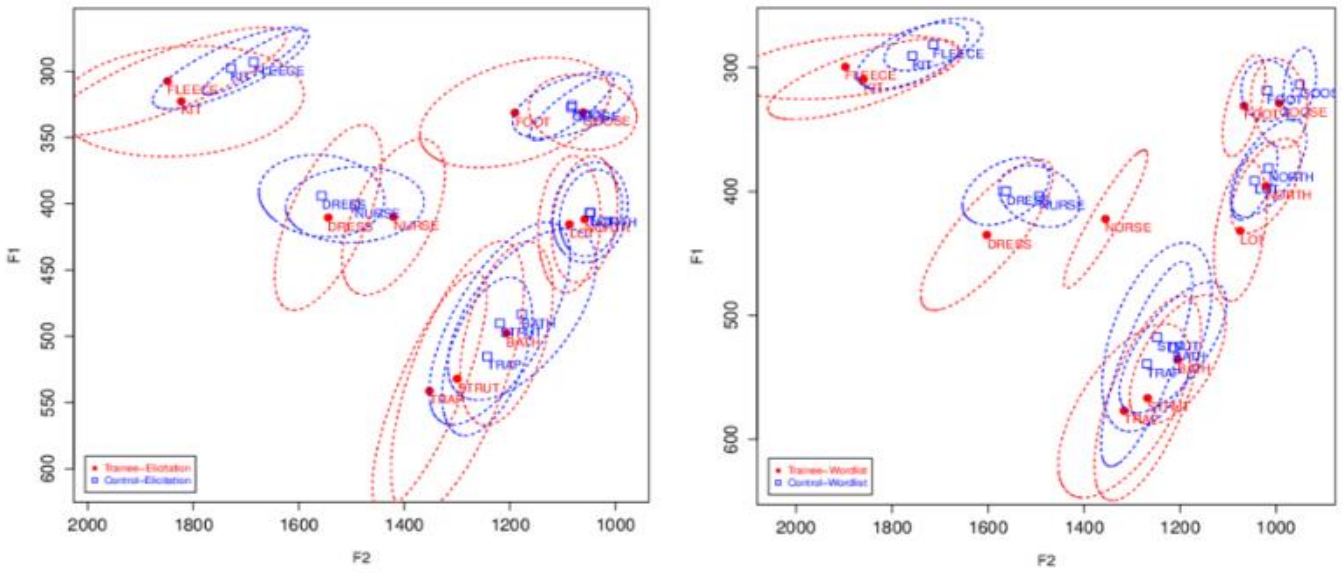


Figure 18. F1xF2 plot of target vowels produced by trainees and controls in the elicitation (left) and wordlist-reading (right) tasks at T2

4.3.3 Comparison of T1 and T2 Spectral Characteristics

Table 17 shows the overlap between T1 and T2 productions of each target vowel produced by trainees and controls in each task (formula: *manova (cbind(F1, F2)~Time, data=data.frame)*). P-values in these analyses are not reported as they were all above the set significance level (0.05). **Figure 19** presents the respective F1xF2 plots, demonstrating the degree of overlap between T1 and T2 values in each group and task. Evidently, the productions of learners remained almost the same at T2, with Pillai scores close to 0 in most cases. The only exceptions with Pillai scores higher than 0.1 were FOOT for trainees in the wordlist-reading task, FLEECE for the control group in the elicitation task, and LOT and FOOT in the control group in the wordlist-reading task. Even in these cases, however, the degree of overlap was still high.

VOWEL	TRAINEES		CONTROLS	
	Elicitation	*Wordlist*	*Elicitation*	*Wordlist*
KIT	0.0073	0.0009	0.0185	0.0229
FLEECE	0.0782	0.0283	0.2298	0.0529
DRESS	0.0273	0.0066	0.0169	0.0607
NURSE	0.0062	0.01	0.0947	0.0155
BATH	0.006	0.0056	0.0394	0.0195
TRAP	0.0014	0.0307	0.0222	0.0443
STRUT	0.0205	0.0319	0.0086	0.0432
LOT	0.0029	0.0167	0.0067	0.1305
NORTH	0.0157	0.0772	0.0914	0.0258
FOOT	0.0485	0.1422	0.0275	0.2563
GOOSE	0.0955	0.0229	0.0955	0.0229

Table 17. Comparison between T1 and T2 productions of each CYG group in each task

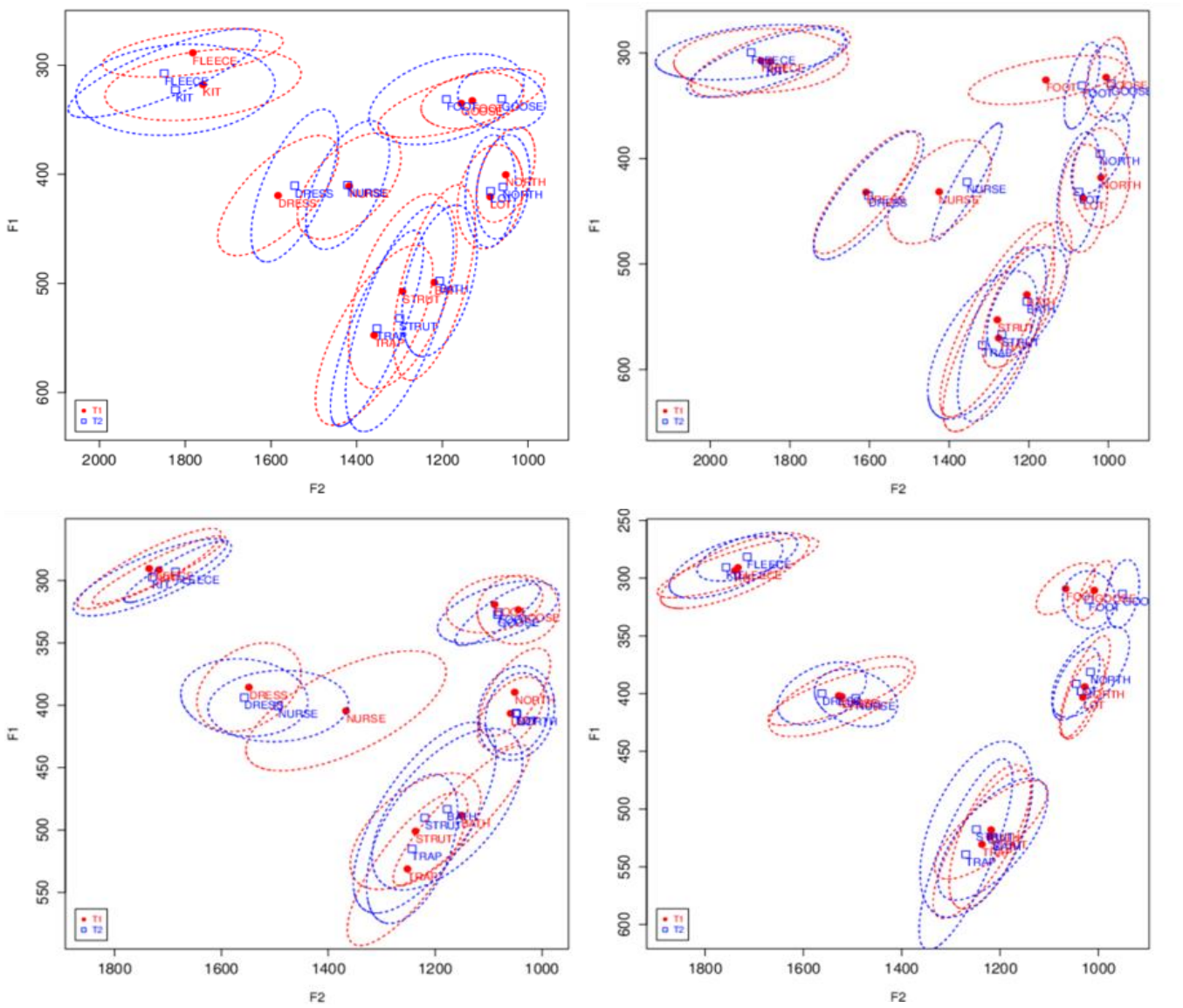


Figure 19. F1xF2 plots of T1 and T2 productions of trainees in the elicitation (top left) and wordlist-reading (top right) task, and of controls in the elicitation (bottom left) and wordlist-reading task (bottom right)

To assess learners' ability to differentiate the vowels of a vowel pair in their productions, the overlap (in Pillai score) between the vowels in a pair was calculated for each task and each group of participants (**Table 18**) (formula: *manova (cbind(F1, F2)~Vowel, data=data.frame)*). Overall, the vowels in a vowel pair tended to be produced with considerable overlap by these learners, although it is worth noting that DRESS-NURSE had the least overlap in the productions of trainees in both tasks. The only other pair with a Pillai score over 0.3 was BATH-TRAP for trainees in the elicitation task and FOOT-GOOSE for both groups in the wordlist-reading task. Comparisons between BATH and LOT also showed a high Pillai score across group and task at T2.

VOWEL PAIR	TRAINEES				CONTROLS			
	Elicitation		*Wordlist*		*Elicitation*		*Wordlist*	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*
KIT-FLEECE	0.0608		0.0714		0.0417		0.0497	
DRESS-NURSE	0.4443	<0.001	0.4709	<0.001	0.1379		0.1031	
BATH-TRAP	0.4047	0.001	0.1908	0.046	0.0984		0.0861	
BATH-STRUT	0.1672		0.0884		0.0306		0.0569	
TRAP-STRUT	0.0773		0.0449		0.0358		0.0203	
LOT-NORTH	0.0571		0.1902		0.0003		0.0538	
FOOT-GOOSE	0.1742		0.3524	0.005	0.0022		0.4093	0.019
BATH-LOT	0.4767	<0.001	0.5325	<0.001	0.4225	0.009	0.7186	<0.001

Table 18. Overlap between vowels of contrastive pairs as produced by trainees and controls in each task at T2

4.3.4 Comparison with NE Speakers

Table 19 lists the mean F1 and F2 values of each vowel produced by NE speakers in the wordlist-reading task in the /bVt/ and /gVt/ contexts, which are plotted in **Figure 20** as well.

VOWEL	F1	F2
KIT	376	1717
FLEECE	297	1973
DRESS	509	1603
NURSE	463	1407
BATH	475	1166
TRAP	548	1409
STRUT	513	1302
LOT	434	1132
NORTH	349	1001
FOOT	364	1421
GOOSE	313	1427

Table 19. Mean Lobanov-normalised F1 and F2 (in Hz) of English vowels produced by NE speakers

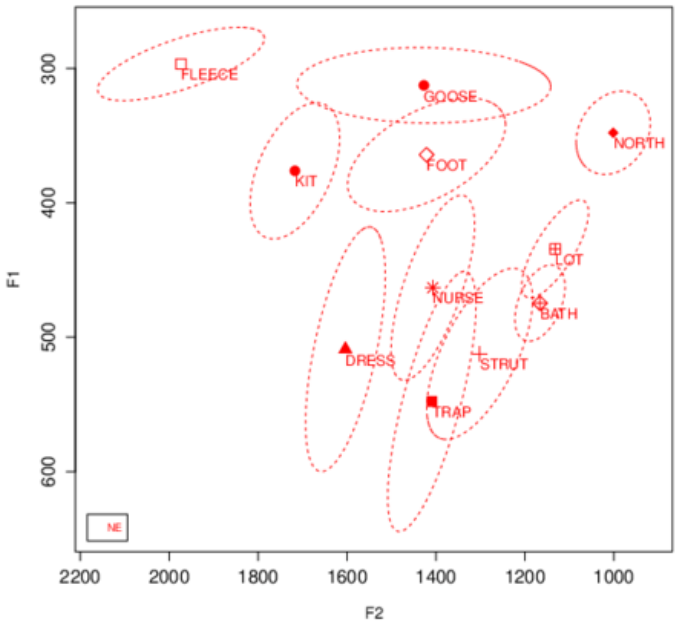


Figure 20. F1xF2 plot of vowels produced by NE speakers

The overlap between vowels in a contrastive pair is shown in **Table 20** (formula: *manova (cbind(F1, F2)~Vowel, data=data.frame)*). As opposed to CYG learners, the vowel productions of NE speakers were more widely spread in the vowel space, and there was less overlap between the members of a contrastive pair.

VOWEL PAIR	PILLAI SCORE	P-VALUE
KIT-FLEECE	0.7765	<0.001
DRESS-NURSE	0.5933	<0.001
BATH-TRAP	0.7304	<0.001
BATH-STRUT	0.3468	0.002
TRAP-STRUT	0.2185	0.013
LOT-NORTH	0.6387	<0.001
FOOT-GOOSE	0.3576	<0.01

Table 20. Overlap between vowels of a vowel pair in NE speakers' productions

The overlap between the productions of NE speakers and CYG learners (both groups) in each task at T1 was calculated using the Pillai score and is shown in **Table 21** (formula: *manova (cbind(F1, F2)~Group, data=data.frame)*). **Figure 21** shows the plotted vowels for each task compared to NE productions. The same analysis was conducted for T2 productions; however, the two CYG groups were analysed separately in this analysis, which therefore involved a smaller sample size. The overlap between NE speakers and each CYG group in each task at T2 is shown in **Table 22** and **Figure 22**.

As evident by these results, KIT, FOOT and GOOSE were the vowels with the least overlap with NE speakers at both T1 and T2, in both tasks and groups. More specifically, KIT was much lower in the productions of NE speakers, whereas FOOT and GOOSE were much fronter, as is common in contemporary SSBE. These were followed by TRAP, which had a high Pillai score at T1 in the wordlist-reading task and in the productions of the controls at T2 in both tasks. On the other hand, BATH and STRUT had a high degree of overlap between learners and NE speakers at both T1 and T2 and in both tasks. However, it should be noted that BATH and STRUT were produced with considerable overlap by all learners at T2, as shown in 4.3.3, and particularly in the case of BATH, the high degree of overlap can be at least partly attributed to the large variation found in learners' productions, which were more widely spread, taking up more space than NE speakers' productions, who produced the vowels with less overlap between them.

VOWEL	ELICITATION		WORDLIST	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT	0.4763	<0.001	0.6248	<0.001
FLEECE	0.3169	<0.001	0.2147	0.006
DRESS	0.4015	<0.001	0.3520	<0.001
NURSE	0.2439	0.01	0.2878	0.002
BATH	0.0295		0.2083	0.012
TRAP	0.2061	0.007	0.5182	<0.001
STRUT	0.0165		0.2176	0.007
LOT	0.1344		0.3167	<0.001
NORTH	0.3027	<0.001	0.3518	<0.001
FOOT	0.5518	<0.001	0.4719	<0.001
GOOSE	0.4315	<0.001	0.5826	<0.001

Table 21. Overlap between NE speakers and CYG learners' (both groups) vowel productions in each task at T1

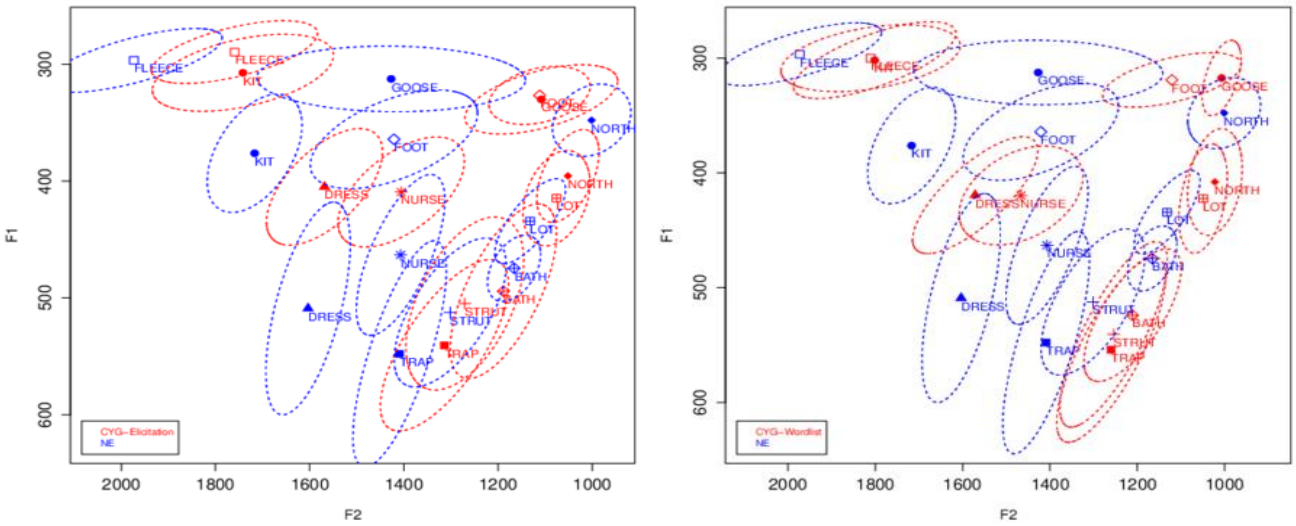


Figure 21. F1xF2 plots of NE speakers and CYG learners' (both groups) productions in the elicitation (left) and wordlist-reading (right) tasks at T1

VOWEL	TRAINEES				CONTROLS			
	Elicitation		*Wordlist*		*Elicitation*		*Wordlist*	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*	*Pillai Score*	*p-value*
KIT	0.3607	<0.001	0.6276	<0.001	0.5461	<0.001	0.6025	<0.001
FLEECE	0.3207	0.002	0.0607		0.5731	<0.001	0.4138	<0.001
DRESS	0.2837	0.009	0.2913	0.007	0.3820	0.004	0.3953	0.002
NURSE	0.2654	0.010	0.1747		0.2952	0.018	0.6123	<0.001
BATH	0.0885		0.3723	0.001	0.0097		0.3029	0.013
TRAP	0.1321		0.3283	0.001	0.5154	<0.001	0.5085	<0.001
STRUT	0.0249		0.2902	0.005	0.0983		0.0865	
LOT	0.1066		0.1998	0.035	0.2528	0.020	0.3254	0.006
NORTH	0.4190	<0.001	0.3337	0.002	0.4418	<0.001	0.1813	
FOOT	0.3094	0.003	0.6277	<0.001	0.5435	<0.001	0.5852	<0.001
GOOSE	**0.4566**	**<0.001**	**0.5444**	**<0.001**	**0.3996**	**0.005**	**0.5554**	**<0.001**

Table 22. Overlap between the vowel productions of NE speakers and each CYG group in each task at T2

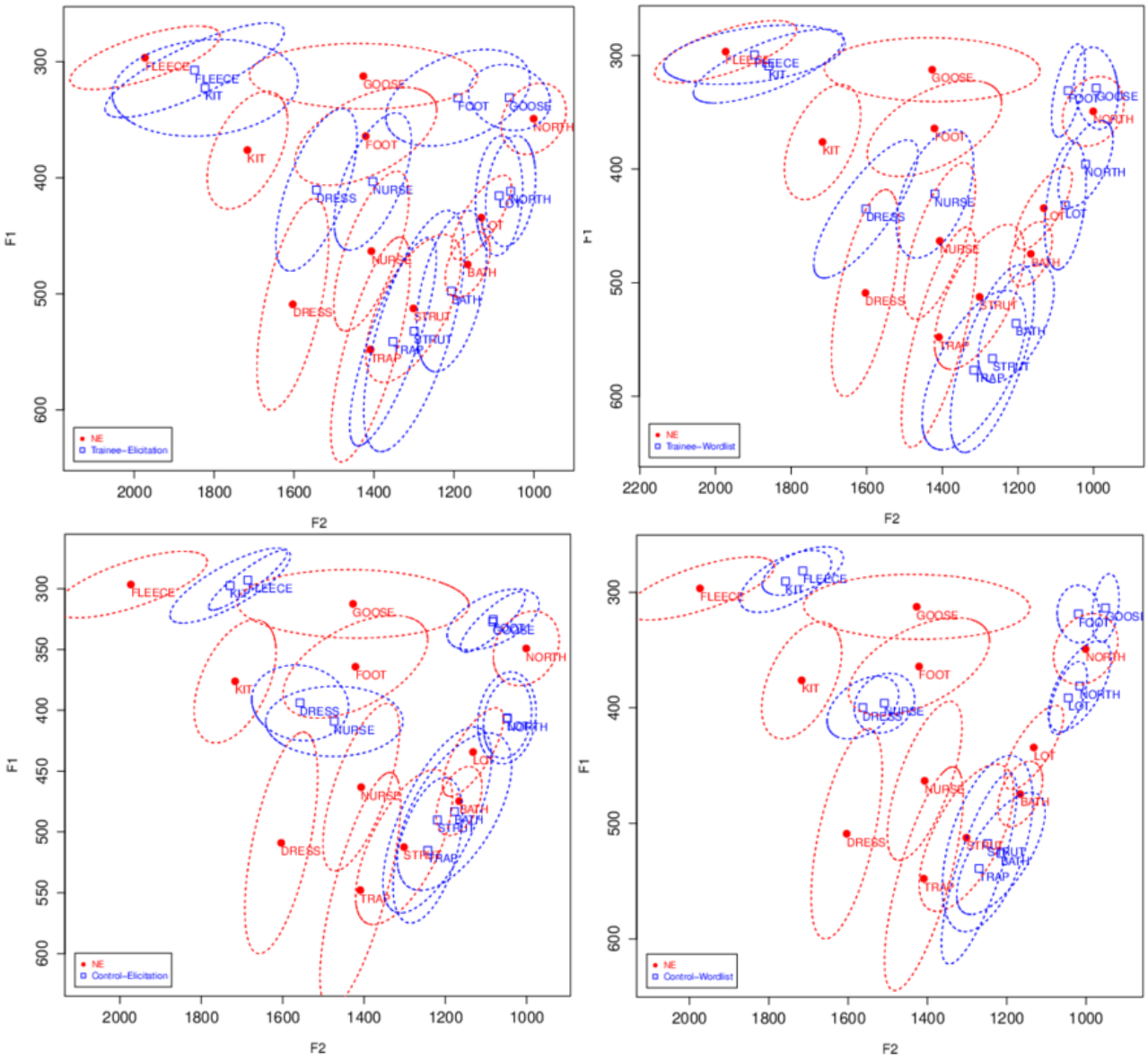


Figure 22. F1xF2 plots of NE speakers' productions compared to the productions of trainees in the elicitation (top left) and wordlist-reading (top right) task and of controls in the elicitation (bottom left) and wordlist-reading (bottom right) task at T2

4.3.5 Comparison with CYG Vowels

The mean Lobanov-normalised formant values for CYG vowels in the /ˈbVtV/ context are given in **Table 23**. A comparison was made between CYG learners' productions of L2 vowels and the corresponding L1 vowel in each of the two tasks (formula: *manova (cbind(F1, F2)~Vowel, data=data.frame)*). The Pillai scores showing the overlap between L1 and L2 vowels as produced by the learners in each task are shown in **Table 24**. **Figure 23** presents the respective F1xF2 plots.

VOWEL	F1	F2
i	309.6	1747.7
e	426.2	1542.7
a	565.1	1281.6
o	419.1	1083.7
u	333.9	1054.2

Table 23. Mean Lobanov-normalised formant values for CYG vowels

VOWELS	ELICITATION		WORDLIST	
	Pillai Score	*p-value*	*Pillai Score*	*p-value*
KIT_i	0.001		0.1014	
FLEECE_i	0.2046	0.016	0.0915	
DRESS_e	0.1309		0.0578	
NURSE_e	0.2613	0.014	0.0859	
BATH_a	0.2574	0.013	0.1346	
TRAP_a	0.1139		0.0099	
STRUT_a	0.1695	0.035	0.0251	
LOT_o	0.0054		0.1073	
NORTH_o	0.1025		0.2104	0.014
FOOT_u	0.1388		0.2757	0.008
GOOSE_u	0.0793		0.2435	0.009

Table 24. Pillai score of the productions of L1 vowels and the corresponding L2 vowel by CYG learners in each task at T1 and significance of result

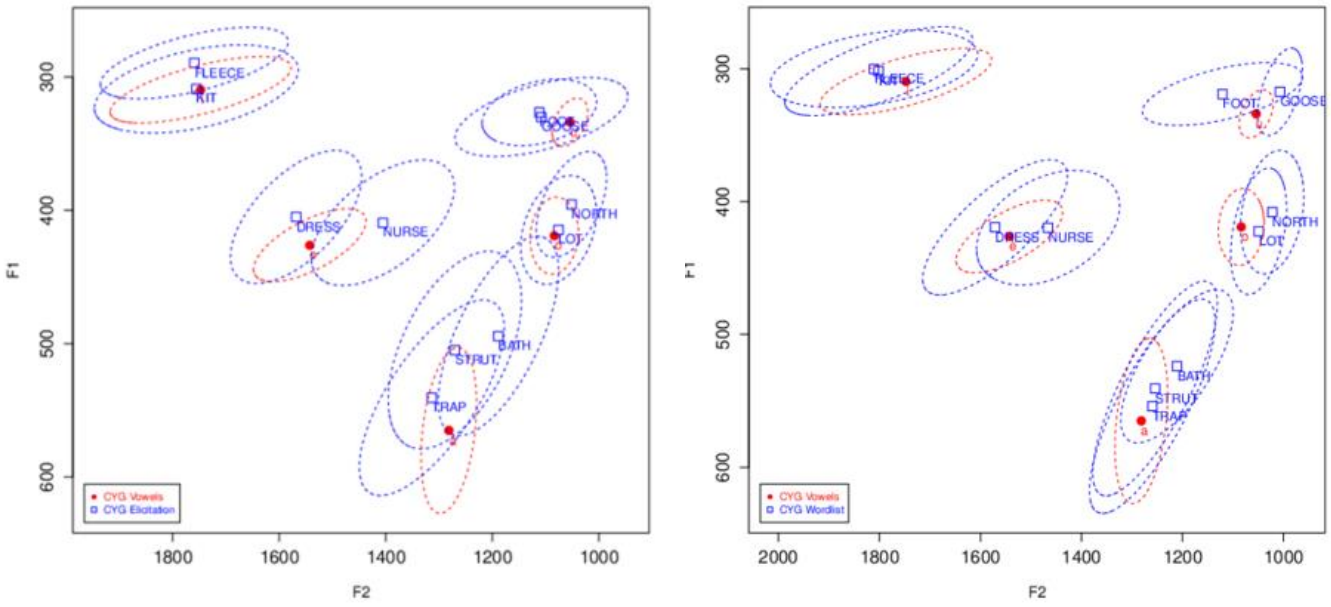


Figure 23. F1xF2 plots of CYG vowels and L2 vowels produced in the elicitation (left) and the wordlist-reading (right) task at T1

As evident from these analyses, CYG learners produced the L2 vowels clustered around their L1 vowels. The most deviation from L1 articulatory routines was observed in the productions of NURSE and BATH in the elicitation task, and in NORTH, FOOT and GOOSE in the wordlist-reading task, although these still had a high degree of overlap with the respective contrastive L2 vowels and the L1 vowels. Due to the small sample size and the lack of

considerable changes between T1 and T2 productions of the learners, this analysis focused on the productions of both CYG groups combined at T1 and was not conducted for T2 data.

Finally, in an attempt to explain the performance of learners in L2 vowel production and their differences with NE speakers, the CYG vowels as produced by CYG speakers and the L2 vowels as produced by the NE speakers were also analysed. The Pillai scores showing the overlap between CYG and English vowels are shown in **Table 25**, and the vowels are plotted in **Figure 24**.

As shown in these analyses, CYG /i/ is found between English FLEECE and KIT, with almost equal distance from each. As opposed to this, CYG /e/ and CYG /o/ are closer to the corresponding short L2 vowels DRESS and LOT, respectively, than the long L2 vowels NURSE and NORTH, respectively. Furthermore, CYG /u/ is almost equally distant from the L2 vowels GOOSE and FOOT, which are considerably fronter, and is closer to the back vowel NORTH. Finally, CYG /a/ has more overlap with L2 STRUT than BATH or TRAP. As will be further discussed in 5.1.2, these may explain learners' deviation from L1 articulatory routines in their productions of NURSE, BATH, NORTH, FOOT and GOOSE as noted above, since their distance from the closest L1 vowels may be sufficient for them to perceive them as different, and therefore attempt to produce them with different spectral characteristics.

VOWELS	PILLAI SCORE	P-VALUE
KIT_i	0.4913	<0.001
FLEECE_i	0.5613	<0.001
DRESS_e	0.276	0.013
NURSE_e	0.5885	<0.001
BATH_a	0.6208	<0.001
TRAP_a	0.557	<0.001
STRUT_a	0.2494	0.016
LOT_o	0.126	
NORTH_o	0.6054	<0.001
FOOT_u	0.6502	<0.001
GOOSE_u	0.5107	<0.001
NORTH_u	0.2427	0.015

Table 25. Overlap between English and corresponding CYG vowels

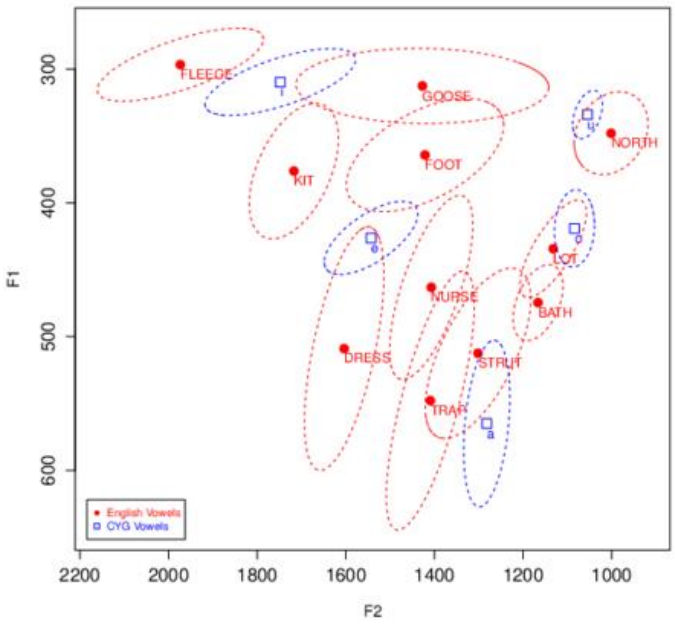


Figure 24. F1xF2 plot of English vowels produced by NE speakers and CYG vowels

Given the lack of a substantial effect of the training on spectral characteristics, no further analyses were conducted to assess the retention of learning at T3, or the generalisation of learning to new contexts.

4.4 Intelligibility Ratings

4.4.1 Individual Participant Performance

Table 26 shows the percentage of correctly identified stimuli produced by each CYG learner in the trained contexts /bVt/ and /gVt/ as per the NE raters' judgments in each task at the 3 time points.

PARTICIPANT	ELICITATION			WORDLIST		
	T1	*T2*	*T3*	*T1*	*T2*	*T3*
CYG01*	41	51	45	48	56	52
CYG03*	63	58	55	71	70	71
CYG04*	56	51	46	55	49	52
CYG05*	55	43	55	55	55	60
CYG06	60	53	62	49	45	45
CYG07	48	40	53	51	56	54
CYG08*	57	61	55	60	65	55
CYG09	48	55	52	45	48	51
CYG10	48	51	49	44	41	42
CYG11	47	39	48	54	49	44
CYG12*	58	52	53	64	58	46
CYG15	54	38	44	53	49	48
CYG16*	44	36	49	40	52	52
CYG17*	55	56	60	53	61	62

Table 26. %-correct responses elicited by NE raters in the intelligibility task for each CYG learner (asterisks represent trainees; underlined values indicate significant changes across time)

The performance of individual participants in each task and at each time point was submitted to a mixed-effects binomial logistic regression analysis, using the formula *glmer (Result~Time+(1|Vowel)+(1|Rater), data=data.frame, family="binomial"*) in order to assess whether each participant performed differently at each time (T1, T2, T3). In the elicitation task, only participants CYG03 (est.=-0.943, p=0.027 at T3), CYG05 (est.=-0.884, p=0.011 at T2) and CYG15 (est.=-0.873, p=0.007 at T2) had a significantly different performance at the three time points, while in the wordlist-reading task, only participants CYG12 (est.=-1.079, p=0.007 at T3) and CYG16 (est.=0.75, p=0.049 at T2; est.=0.819, p=0.038 at T3) had a significantly different performance at different time points. Of these, only participant CYG16 in the

wordlist-reading task showed improvement, suggesting that the training may have been beneficial for this participant but only in this task. Participants CYG03, CYG05, CYG15 and CYG12, three trainees and one control, demonstrated a lower performance at T2 and/or T3 compared to their T1 performance.

Task effects were also examined using the formula *glmer (Result~Task+(1|Vowel)+(1|Rater), data=data.frame, family="binomial"*) on each participant's performance at T1. No task effect was found for any of the participants, indicating that they performed similarly in the two tasks. Finally, correlation analysis was conducted (formula: *cor.test(data$x, data$y, method = 'pearson')*) to examine whether individual learners' performance in the perceptual test and their intelligibility scores at T1 were correlated. This analysis showed no correlation between perception and either of the production tasks. T2 and T3 correlations were not examined, due to the small sample size that would emerge by splitting the two CYG groups into trainees and controls.

As will be further discussed in 5.2.3.2, based on the fact that other participants, including both trainees and controls, had a lower performance at T2 and/or T3 compared to T1 as well, albeit insignificant in other cases, it is hypothesised that participants may have been overwhelmed or confused upon becoming aware of the existence of vowel contrasts in English. This may have led to an unsuccessful attempt to identify the appropriate cues to differentiate between the members of the contrast, whether learners belonged to the experimental or the control group.

4.4.2 Group Performance

Figure 25 shows the production performance of the two groups in each task and time in trained contexts as rated by NE raters in the intelligibility task. In the examination of whether participants performed differently in the two tasks at different time points (formula: *glmer (Result~Task+(1|Subject)+(1|Vowel)+(1|Rater), data=data.frame, family="binomial")*), the results showed a significant difference between the two tasks only at T2 (est.=0.288, $p<0.001$), with productions being rated as more intelligible in the wordlist-reading task. Further analysis to examine whether there was a difference between trainees and controls in each task at each time point (formula: *glmer (Result~Task*Group+(1|Subject)+(1|Vowel)+(1|Rater), data=data.frame, family="binomial")*) demonstrated a significant effect of the interaction at T3 (est.=0.443, $p=0.012$) only, suggesting that the performance of the two groups was similar

in the elicitation task but significantly different in the wordlist-reading task, in which case trainees outperformed controls.

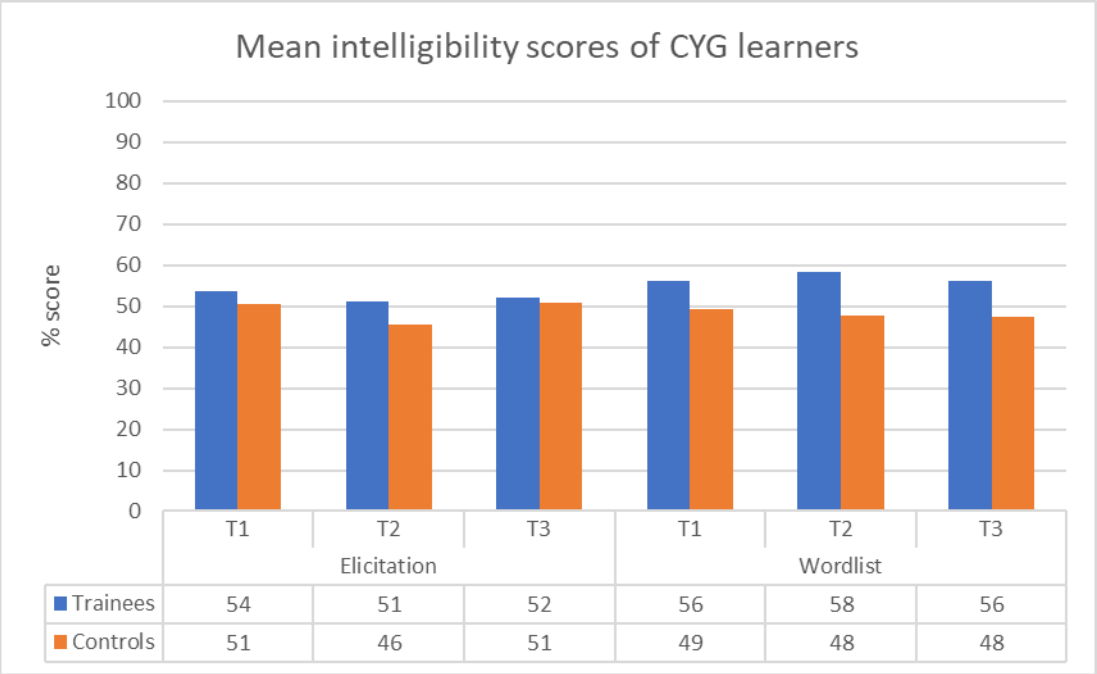


Figure 25. Mean intelligibility scores of the two CYG groups in each task and at each time point

The two groups were further compared between them at each time point using the formula *glmer (Result~Group+(1|Subject)+(1|Vowel)+(1|Rater), data=data.frame, family="binomial")*. The overall analysis of both elicitation and wordlist-reading data combined showed that, whereas the two groups performed similarly at T1, they had significant differences at T2 (est.=0.409, p=0.009) and at T3 (est.=0.273, p=0.046), with trainees outperforming controls in both cases. Further analyses were conducted to examine whether this difference was in the elicitation task, the wordlist-reading task or both, using the same formula on different subsets of the data frame. This analysis showed no significant differences between the two groups in the elicitation task at any time point, or in the wordlist-reading task at T1. Significant differences were only found in the wordlist-reading task at T2 (est.=0.587, p<0.001) and T3 (est.=0.487, p=0.008).

Each group was also examined to determine whether they performed differently as an effect of *Time* (formula: *glmer (Result~Time+(1|Subject)+(1|Vowel)+(1|Rater), data=data.frame, family="binomial")*). No significant effect of *Time* was found in the performance of either trainees or controls when the data for both tasks were combined.

Further examination on elicitation and wordlist-reading data separately, showed that the only difference that reached significance was in the performance of controls at different time points in the elicitation task (T1 vs. T2: est.=0.299, p=0.03; T2 vs. T3: est.=0.297, p=0.03). Therefore, it can be concluded that the training was not adequate to cause any changes in the overall performance of trainees among the three time points.

4.4.3 Individual Vowel Performance

Table 27 demonstrates the percentage of correctly identified target vowels in the known contexts /bVt/ and /gVt/ by NE raters for trainees and controls in each task and time. Overall, the most intelligible vowels produced by CYG learners according to the ratings were FLEECE, and the members of the DRESS-NURSE contrast, while the most problematic vowels were STRUT and KIT, followed by BATH and NORTH in most cases.

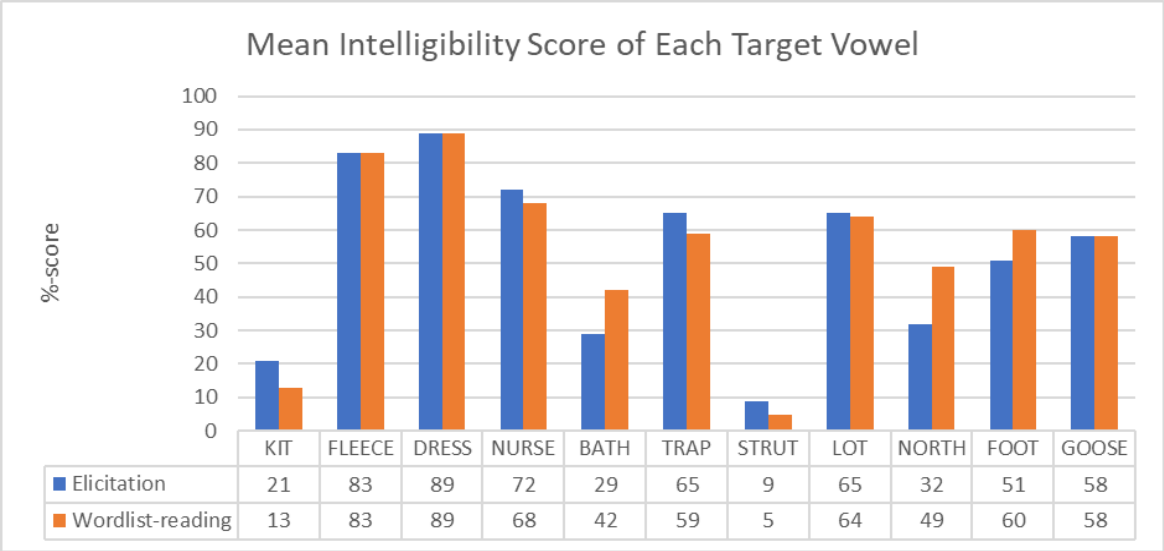
VOWEL	**TRAINEES**						**CONTROLS**					
	Elicitation			*Wordlist*			*Elicitation*			*Wordlist*		
	T1	*T2*	*T3*	*T1*	*T2*	*T3*	*T1*	*T2*	*T3*	*T1*	*T2*	*T3*
KIT	21	27	22	16	12	20	22	17	24	8	15	16
FLEECE	83	**66**	88	80	**89**	87	83	73	80	87	85	78
DRESS	88	76	84	90	85	81	90	85	83	87	87	92
NURSE	72	77	77	90	95	85	70	60	50	35	29	25
BATH	24	**45**	25	53	63	63	35	28	22	30	30	30
TRAP	69	63	63	59	64	56	60	50	65	58	56	58
STRUT	9	13	4	1	4	4	8	2	10	11	5	5
LOT	72	**55**	63	65	64	60	56	65	73	62	60	57
NORTH	32	35	34	51	55	52	32	22	24	47	36	52
FOOT	53	46	56	60	57	62	50	48	53	60	**83**	47
GOOSE	62	59	50	60	63	60	53	42	60	56	44	53

Table 27. Intelligibility scores obtained for each target vowel produced by trainees and controls in each task and time (values in bold indicate significant changes across time)

To examine the effect of *Time* (T1, T2, T3) on each vowel in each task and by each group, different subsets were created and the formula *glmer (Result~Time+(1|Subject)+(1|Rater), data=data.frame, family="binomial")* was run. In trainees, a significant effect of *Time* was observed in FLEECE (T2: est.=-1.266, p=0.005), BATH (T2: est.=1.059, p=0.016) and LOT (T2: est.=-0.843, p=0.027) in the elicitation task and in FLEECE (T2: est.=1.023, p=0.042) in the wordlist-reading task. However, this improvement was only positive in BATH, which had a higher intelligibility at T2 but this was not retained at T3, and in FLEECE in the wordlist-reading

task, where the intelligibility score at T3 remained higher than at T1, but without reaching significance. LOT became less intelligible at T2, and FLEECE in the elicitation task became less intelligible at T2 before it became more intelligible at T3 compared to T1. In the control group, the effect of *Time* reached significance only in FOOT in the wordlist-reading task (T2: est.=1.231, p=0.031), which was more intelligible at T2 but less intelligible at T3 compared to T1.

Further analyses to examine the effects of *Task* and *Vowel* on the intelligibility of target vowels at T1 were conducted to assess how the productions of both CYG learners combined were perceived by NE raters, in order to complement the results of acoustic analyses. Since improvements were limited to very few target vowels, the effects of *Task* and *Vowel* at T2 and T3 were not conducted. **Figure 26** shows the mean %-correct identification scores of the target vowels as rated by NE raters for both CYG groups at T1 in each task for trained contexts.



	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
Elicitation	21	83	89	72	29	65	9	65	32	51	58
Wordlist-reading	13	83	89	68	42	59	5	64	49	60	58

Figure 26. Mean intelligibility scores of each target vowels for CYG trainees and controls combined at T1 in each task

The effect of *Task* was examined using binomial logistic regression (formula: *glmer (Result~Task+(1|Subject)+(1|Rater), data=data.frame, family="binomial")*) and reached significance in only three vowels: KIT (est.=-0.9916, p=0.0145), which was found to be more intelligible in the elicitation task, and BATH (est.=0.7832, p=0.0185) and NORTH (est.=0.9082, p=0.00162), which were found to be more intelligible in the wordlist-reading task. The effect of *Vowel,* examined using binomial logistic regression (formula: *glmer (Result~Vowel+(1|Subject)+(1|Rater), data=data.frame, family="binomial")*), reached significance in all vowel pairs examined in both tasks, with the exception of the FOOT-GOOSE

contrast, as shown in **Table 28**. This demonstrates that one of the vowels in each of these pairs was produced as significantly more intelligible than the other.

VOWEL PAIR	ELICITATION		WORDLIST	
	Estimate	*p-value*	*Estimate*	*p-value*
KIT-FLEECE	-3.006	<0.001	-4.101	<0.001
DRESS-NURSE	-1.203	<0.001	-1.795	<0.001
BATH-TRAP	2.065	<0.001	0.962	0.001
BATH-STRUT	-1.443	<0.001	-2.608	<0.001
TRAP-STRUT	3.093	<0.001	3.816	<0.001
LOT-NORTH	-1.467	<0.001	-0.593	0.018
FOOT-GOOSE	0.27	0.3	-0.066	0.806

Table 28. Results of binomial logistic regression on the effect of *Vowel* on intelligibility score

Table 29 and **Table 30** present the confusion matrices with NE raters' percentage of responses to each target vowel and the most commonly confused vowels as produced by all CYG learners at T1 in the elicitation and wordlist-reading tasks, respectively (confusion matrices for the two groups and two tasks separately at T1 and T2 are available in Appendix D). It is evident from these that the predicted confusion patterns are for the most part confirmed by the responses of the raters. One noteworthy exception is the case of TRAP, which was expected to be mostly confused with BATH and/or STRUT, but was instead mostly perceived as LOT by NE raters in both tasks. The same can be observed for BATH, although in the wordlist-reading task, this vowel was better perceived as the correct vowel, followed by the expected TRAP and then LOT.

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**21.4**	**69.3**	8.6								
FLEECE	**12.6**	**83**	1.5								2.2
DRESS	1.4		**88.6**	4.3				1.4			
NURSE	2.4		**15.3**	**71.8**	1.2		1.2	1.2		4.7	2.4
BATH				5.6	**28.9**	14.4	7.8	**38.9**	4.4		
TRAP			2.2	3.7	9.6	**65.2**	5.2	**14.1**			
STRUT			3.2	5.6	**11.2**	**45.6**	8.8	24	1.6		
LOT							5.4	**65.4**	**23.1**	2.3	2.3
NORTH							3.7	**57**	**31.9**	2.2	4.4
FOOT				1.4				1.4	1.4	**51.4**	**43.6**
GOOSE				1			1	1.9	3.8	**34.3**	**58.1**

Table 29. Confusion matrix of NE raters' percentage of responses to CYG learners' productions of each target vowel at T1 in the elicitation task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**12.9**	**85**									1.4
FLEECE	**11.5**	**83.1**	2.3		1.5						1.5
DRESS			**88.6**	7.9						1.4	
NURSE			**28**	**68**	1	1		2			
BATH				3.1	**42.3**	**25.4**	3.1	24.6	1.5		
TRAP					13.3	**58.5**	3	**23**	1.5		
STRUT					10.8	**54.6**	5.4	**26.2**	1.5		
LOT				1.5		1.5		**63.7**	**30.4**		
NORTH							1.5	**42.3**	**49.2**	2.3	1.5
FOOT				1.9						**60**	**38.1**
GOOSE									4.8	**35.2**	**58.4**

Table 30. Confusion matrix of NE raters' percentage of responses to CYG learners' productions of each target vowel at T1 in the wordlist-reading task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

4.4.4 Generalisation and Retention

Figure 27 presents the overall %-correct identification of stimuli produced by each group in the untrained contexts /sVt/ and /dVt/ in each task at the three time points. In order to assess the generalisation of learning to new contexts and retention at T3, the effects of *Time* (T1, T2, T3), *Context* (known, unknown), *Group* (trainees, controls) and *Task* (elicitation, wordlist-reading), were examined running separate mixed-effects binomial logistic regression models

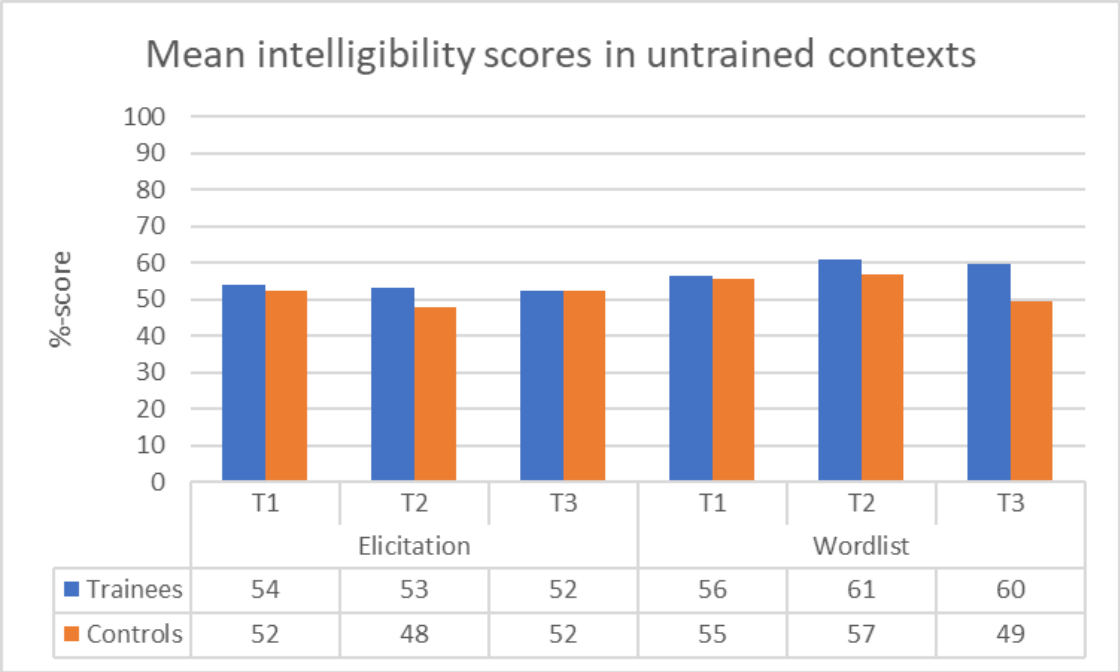


Figure 27. Mean intelligibility scores of the two CYG groups in untrained contexts in each task at T1, T2 and T3

with each IV as the fixed effect and *Subject*, *Vowel* and *Rater* as the random effects (formula: glmer*(Result~IV+(1|Subject)+(1|Vowel)+(1|Rater), data=data.frame, family="binomial"))*.

The effect of *Time* in the analysis of stimuli in new contexts did not yield any significant results for the experimental group in either task, although it did for the control group, who had a significantly lower intelligibility at T3 compared to T1 productions in the wordlist-reading task (est.=-0.339, p=0.046). In addition, trainees' intelligibility scores were not affected by *Context* either, since no significant effect of this variable was found. The only comparison that yielded a significant result in this analysis was that of the performance of controls in the wordlist-reading task at T2, where the learners performed significantly worse in the /bVt/ and /gVt/ contexts than in the /sVt/ and /dVt/ contexts. The effect of *Group* was also limited; the two groups performed similarly between them across task and time in new contexts, except in the wordlist-reading task at T3, where trainees had significantly more intelligible productions (est.=0.601, p=0.013) according to NE raters. Finally, the analysis of the effect of *Task* reached significance in more comparisons, with wordlist-reading productions being rated as more intelligible than productions in the elicitation task. This difference reached significance for trainees at T2 (est.=0.391, p=0.008) and T3 (est.=0.374, p=0.014) and for controls at T2 (est.=0.385, p=0.03).

4.5 Correlation Analyses

In order to examine the effects of individual differences in motivation, input and L1-L2 use patterns on participants' perceptual and production performance, two-tailed Pearson correlations were conducted in R using the general formula *cor.test(data$x, data$y, method = 'pearson')*. Perceptual performance included three DVs, one for each time point, and it was measured as each participant's %-correct identification score in the known contexts /bVt/ and /gVt/. Production performance was measured through the overall intelligibility score of each participant, i.e. the percentage of their productions that was correctly identified by the 5 NE raters in the intelligibility task in their productions in known contexts. Six separate DVs were included, one for each task (elicitation, wordlist-reading) at each time point (T1, T2, T3). Before running the analyses, it was firstly established that the data were normally distributed, using the Shapiro-Wilk test. All p-values were greater than the significance level 0.05, and therefore, normality was assumed.

4.5.1 Motivation

Questionnaire items 34, 35, 38, 39, 40 and 42 (reiterated in **Table 31**) were used to measure each participant's motivation score. On the 6-point-scale questions, participants received a score from 1 to 6 for each of the items 34, 35, 38 and 42. They received a score of 1 (pronunciation is not important), 2 (pronunciation somewhat important or not specified), or 3 (pronunciation is important) for their responses on items 39 and 40, which was then multiplied by two to ensure that answers to these items were weighted equally with 6-point-scale questions. The total score of motivation was then calculated for each participant and was used in correlation analyses with the overall intelligibility and perceptual score of each participant.

NO.	QUESTION
34.	In your opinion, how important is it to learn English?
35.	How important do you consider a good pronunciation of English to be?
38.	In the past, I have practised my pronunciation skills a) on my own, b) while watching videos, c) while watching movies or TV, d) while listening to recordings for pronunciation improvement, e) by observing the speech of English-speaking relatives or friends, f) by asking someone to help me practise, g) through pronunciation training, h) in any other way.
39.	When communicating in English, it is more important for people to understand what I want to say/to have a good or native-like accent/I do not think my accent is important/other (select all that apply).
40.	Do you think having a good English pronunciation has any positive effects in any aspect of your professional or personal life?
42.	How much do you try to improve your English pronunciation?

Table 31. Questionnaire items used for correlation analyses for motivation

Participants' motivation score was positively correlated with their intelligibility score in the elicitation ($R=0.666$, $p=0.009$) and wordlist-reading ($R=0.828$, $p<0.001$) task at T1, but not with any task at T2 or T3. Motivation was not significantly correlated with participants' perceptual performance at any time.

4.5.2 Input and Language Use Patterns

Questionnaire items 27 and 31 (reiterated in **Table 32**) were used to calculate a total score of input and L1-L2 use patterns during participants' school years and university years, respectively. On the 6-point scale, participants received a score from 1 to 6 depending on their response on each sub-item (Never-All the time). The scores for each sub-item were added to obtain a total score of input and L1-L2 use patterns during school years and a total score of

input and L1-L2 use patterns during university years. An overall score was also calculated combining both categories. Sub-items d) and e) in questionnaire item 30 (reiterated in **Table 32**) were used to assess the amount of input and L1-L2 use patterns in participants' EFL classroom. A score from 1 to 6 was assigned depending on participants' responses (Never-All the time). These scores were then used in correlation analyses with the overall intelligibility and perceptual score of each participant.

NO.	QUESTION
27.	When you were at middle school and high school, how often did you use (spoke or listened to) English: a) in the English classroom, b) in other classes, c) at home, d) in social settings?
30.	To what extent were your English classes at school conducted d) in English, e) in Greek?
31.	How often do you use (speak or listen to) English during your university years: a) in class, b) at home, c) in social settings?

Table 32. Questionnaire items used for correlation analyses for input and language use patterns

Whereas no significant correlations were found between school, university or total input on the one hand, and intelligibility score on the other, total input was positively correlated with participants' perception scores at T1 (R=0.562, p=0.036). Interestingly university and total input were found to be positively correlated with motivation as measured in the previous sub-section (R=0.612, p=0.02 and R=0.541, p=0.046, respectively). Furthermore, school and university input were positively correlated (R=0.61, p=0.021), indicating that learners who had or sought opportunities to use the L2 during the early years of L2 instruction, continued to seek such opportunities in their adult life.

4.6 Qualitative Analysis

Following quantitative analyses, trainees CYG01, CYG03 and CYG08 were invited to interviews to obtain more information about their motivation, language use patterns and input received through their experiences. CYG01 was invited as a representative of participants with lower intelligibility scores at T1, who demonstrated numerical improvement at T2. CYG16 was also invited for the same reason, but did not respond to the invitation. CYG03 and CYG08 were invited as representatives of two of the highest scores at T1. Transcripts of the three interviews can be seen in Appendix E.

4.6.1 Motivation

The qualitative analysis of the interviews revealed some important patterns in terms of motivation. More specifically, the two participants who had high intelligibility scores at T1 (CYG03 and CYG08), both reported that they were more motivated to actively try and improve their accents, viewing the training as an opportunity to learn. They mentioned that they had previously made attempts to improve their pronunciation (e.g. listening to native speakers and trying to pronounce words in the same way in front of a mirror), and that they would like to have a "nice" accent as a personal goal (CYG08: "But personally, ideally, I would like to have a nice accent"). CYG01 on the other hand, reported a neutral stance towards improving his pronunciation, stating that pronunciation is important for self-improvement, that "your pronunciation should be correct, to speak English correctly" and that "when you speak correctly, with a correct pronunciation, it's even better for jobs as well", but reported no willingness or motivation to make any effort to improve it, and stated that he had never done anything that requires any effort to improve his pronunciation. However, he mentioned that he was interested in any activities that do not require much effort, such as watching movies, which he felt had helped him improve with minimal conscious effort.

In terms of future goals and their beliefs about employers or lecturers' expectations, all three participants believed that their pronunciation is good enough to meet the expectations of future employers or lecturers, that speaking "correctly" or having a good pronunciation can have a positive effect in professional life, but that their pronunciation would not cost them any career opportunities. CYG01 and CYG08 both expressed the importance of being understandable and expressing themselves clearly rather than being native-like and that having a good or native-like pronunciation is only a matter of self-improvement, which comes at a point where one masters the language, as an optional final step. The difference between the two participants in this respect lay in their motivation to try to improve their pronunciation, as stated above (CYG08: "For me, at this stage, the most important thing is to communicate, to be understood and to understand the other person and to be able to communicate by expressing what you want to say… But personally, ideally, I would like to have a nice accent").

Furthermore, CYG08 also expressed her concerns about stereotyping, i.e. that her clients might consider her less competent in her job due to her accent, although she believed

that “when you are good at your job, sounding native-like can be overcome”. At the same time, CYG03 expressed the most positive attitudes towards improving her pronunciation and her willingness to take any opportunity to improve, for self-improvement purposes, because her pronunciation was not at the level she would have wanted it to be (“because I’m still not talking with the same way that as I want”), but also because she felt that it can help her achieve her goal of studying abroad as well as opening up opportunities “not only for travelling, and for jobs and for your future and so many things”.

4.6.2 Input and Language Use Patterns

As concerns input and language use patterns, all three interviewees reported having family that permanently resides in an English-speaking country; however, they differed in the amount of time spent with them and their language use patterns when meeting them. CYG01 reported very little contact with family members from abroad (one or two days per year), during which they speak in Greek. CYG08 stated that she has relatives in England, the USA and Australia, some of which have returned to Cyprus permanently, while others still living abroad. The interviewee reported more frequent communication with family who has returned, with whom she stated speaking English about 50% of the time or less (“maybe more in Greek, or a mix between the two”), while she mentioned that she speaks mostly in English and uses Greek a little with relatives who only visit once a year for approximately one week. Finally, CYG03 reported speaking English with her visiting relatives, who visit approximately once every two years for about two to three weeks, and only using Greek when it is difficult to express herself in English. Furthermore, she reported visiting them in England three to four times per year, for periods ranging from two weeks to one month each time, as opposed to CYG01 and CYG08 who had never visited an English-speaking country.

In terms of their school experience, all interviewees reported similar experiences, i.e. that most of the work was done at the institutes rather than at school, that their teachers were Greek-Cypriots and that they only had limited opportunities to practice their pronunciation. CYG08 reported that they used mostly Greek in class, except in the final stages when they used English more, but still used Greek in class. CYG01 reported using English in class about 50% of the time, and CYG03 reported speaking English in class most of the time. Most activities related to pronunciation mentioned were speaking activities, in which participants recalled their teachers correcting students’ pronunciation in some cases by

repeating what was said using their own pronunciation. CYG08 noted that teachers focused mostly on grammar, and that listening and speaking activities were included mostly in the final stages, with very little attention paid to pronunciation. Both CYG01 and CYG08 reported that teachers paid attention to pronunciation only during speaking activities, while CYG01 also added that they used to watch Ted Talks in class with the instruction to focus on pronunciation.

At university, the interviewees mentioned that all or almost all their lectures were in English and that they also had at least one module focusing on their English language skills. CYG01 and CYG03 reported only speaking English in class, while CYG08 reported mostly speaking English in class. They mentioned that all or most of their lecturers were Greek-Cypriots or other non-native speakers of English, although CYG03 and CYG08 mentioned having a few lecturers with native-like pronunciations and CYG08 also stated that she had one native English teacher. However, interviewees were unable to specify the accent of these lecturers, possibly because, as CYG08 stated, "they are trying to bring [their accent] down a little so that it's understood". CYG03 and CYG08 also mentioned using English socially during their university years. More specifically, CYG03 reported hanging out with speakers of other languages about two times a week, comfortably communicating in English with them, and CYG08 reported that she had been using English as much as Greek during her university years and that she had joined several Erasmus programmes, travelling to countries that required speaking English with speakers of other languages.

As concerns other activities which involve English language use, interviewees mentioned watching movies, series, videos or podcasts to varying degrees; CYG01 reported many hours of movie-watching per day, at least when younger, CYG03 reported watching movies about once a day, but less frequently at present, and CYG08 mentioned that she used to do it but not very frequently. CYG01 also mentioned working at a job requiring him to speak English with foreigners only.

As evident from the interviews, participants' input was mostly from non-native speakers of English, including their Greek-Cypriot teachers and classmates at school and university, although a few of these teachers were described as native-like. Most native-like input derived from watching movies or other forms of audiovisual input, and interviewees reported that the input received was from various varieties, described broadly as British and American. It should be noted that CYG08 mentioned almost no input received from British speakers, while CYG03

reported extensive use of English with British speakers due to her visits in England and her communication with English-speaking friends and family.

Finally, participants also expressed their preferences in using each language. CYG01 stated that he prefers speaking Greek with anyone who can understand it, and although he stated that he is comfortable using English, he only does so if necessary for communication. On the other hand, CYG03 and CYG08 mentioned that they seek out opportunities to use the English language, but stated that while comfortable using English, their confidence sometimes depends on the topic of conversation and the interlocutors. For instance, CYG03 mentioned that "sometimes I don't feel comfortable to express myself in English and due from our foreigners that we have that they are from Erasmus" and that "actually on the lectures on data, when it's more… economics and accounting basis which are more numerical things, I'm not expressing myself, but on the more theoretical lectures, I can express myself […] Actually, I can partly express myself". Similarly, CYG08 stated that "I feel more confident [speaking English] in the field of psychology, where I'm more familiar with the vocabulary and the expressions from my studies. There are aspects in everyday life, because I'm in Cyprus, and there is vocabulary I'm not familiar with in more everyday things, and it may be a little challenging for me and make me feel somewhat less confident…" and that "When there are people that have a really good accent, I feel somewhat… less confident. When there are people for example from other countries, who also don't have the best accent, it doesn't bother me at all".

4.7 Summary

This chapter has presented the findings of the study in relation to the perceptual performance of learners in each group (4.1.1), individually (4.1.2), in each target vowel (4.1.3), and in tests with new speakers and contexts (4.1.4). These findings showed that the two groups' overall performance was different only at T3, and that both groups improved from T1 to T2 and from T1 to T3. In individual performance, it was observed that most trainees and two controls improved across time, while one trainee had a lower performance at T2 compared to T1. Participants' perception of contrastive vowels confirmed predictions in most but not all cases, and the training seemed to be effective in the identification of individual vowels, since trainees' performance improved across time for most vowels, while controls only improved in KIT. In the generalisation tests, it was found that trainees performed better in known than in

unknown contexts, but the lack of any other statistically significant differences indicated that participants' performance was largely unaffected by the inclusion of untrained contexts or new speakers.

Acoustic analyses of duration and spectral characteristics were detailed in 4.2 and 4.3, respectively, and involved the analysis of T1 and T2 productions, comparisons of productions across time and between CYG learners and NE speakers, comparisons of L1 and L2 vowels and generalisation and retention of learning. In terms of duration, there was a tendency for trainees but not controls to modify the duration of some vowels (STRUT, KIT, DRESS and LOT) towards NE patterns across time, although this was not consistent across tasks. Furthermore, as opposed to NE speakers, CYG learners did not consistently maintain duration differences between the members of a contrastive pair, which reflected the fact that CYG vowels were not differentiated by duration. Furthermore, CYG vowels were found to be shorter than L2 long vowels and approximated the durations of the corresponding L2 short vowels in each set.

Spectral analyses showed a large amount of overlap in the production of members of contrastive pairs, which corresponded to the position of L1 vowels in the vowel space. This was contrary to NE speakers' productions, which made use of a wider area of the vowel space. In addition, there was a high degree of overlap in learners' productions across group, task and time, suggesting that the training was not sufficient for trainees to alter their articulatory configurations for the target vowels. Finally, CYG and English vowels as produced by the respective native speakers were plotted to demonstrate their relative positions and explain possible difficulties for L2 learners.

The findings of the intelligibility task presented in 4.4 showed that only one trainee improved significantly in intelligibility, while three trainees and one control demonstrated a lower performance at T2 and/or T3 compared to T1. While the two groups performed similarly at T1, trainees outperformed controls in intelligibility ratings at T2 and T3; however, neither group improved significantly across time. Overall, the most intelligible vowels produced by CYG learners were FLEECE, and the members of the DRESS-NURSE contrast, while the most problematic vowels were STRUT and KIT, followed by BATH and NORTH in most cases. NE raters' confusion patterns confirmed the predicted patterns for most vowels. Improvement in individual vowels was only observed in two vowels for trainees (BATH in the elicitation task and FLEECE in the wordlist-reading task) and one for controls (FOOT in the wordlist-reading

task). In the absence of substantial improvements in the experimental group, generalisation and retention analyses were limited, and showed no effect of time or context for trainees. The two groups performed similarly across task and time, except in wordlist-reading at T3, when trainees performed better than controls, and both groups produced more intelligible vowels in the wordlist-reading than the elicitation task.

Finally, sections 4.5 and 4.6 addressed the issue of individual differences in participants' perceptual and production performance. The results demonstrated that motivation was positively correlated with intelligibility at T1, but not with perceptual performance, while the total input received was positively correlated with perceptual performance at T1. Furthermore, university and total input received were found to be positively correlated with motivation. This analysis also suggested that learners who had or sought opportunities to use the L2 during the early years of L2 instruction, continued to seek such opportunities in their adult life, as shown by the correlation between school and university input received. These results were further supported by the qualitative analysis presented in 4.6.

The following chapter discusses these findings in relation to each RQ and previous research. It begins with a detailed examination and discussion of the perception and production of the target L2 English vowels by CYG learners and a comparison with NE speakers, and then proceeds to the discussion of the effects of the HVPT paradigm followed in the study. The findings concerning the generalisation and retention of learning are also examined, and finally, the role of individual differences is discussed.

CHAPTER 5: DISCUSSION

5.1 Perception and Production of English vowels by CYG learners

The first part of the discussion aims to answer RQs 1 “How do CYG adult learners perceive and produce L2 English vowels?” and 2 “What are the spectral and durational differences in English vowel production between CYG learners and NE speakers?”, and therefore focuses on the data collected at T1. The study’s overall findings are summarised and discussed first, followed by a detailed description of the learners’ perception and production of contrastive L2 vowels. Finally, a general discussion of the findings in relation to these RQs is provided.

5.1.1 Overall Findings

In terms of perception, participants’ low percentage of correct identification scores at pre-test (41.1% for trainees, 35.7% for controls) demonstrates that L2 English vowel perception is challenging for these learners. In the perception of the vowels in a contrast, CYG learners showed difficulties in perceiving both members of the KIT-FLEECE contrast, better identification of DRESS than NURSE, better identification of TRAP than BATH or STRUT, and better identification of LOT than NORTH. FOOT and GOOSE were identified at a similar rate, higher than all other vowels.

Duration analyses showed that even though this is a secondary cue in vowel production and perception in English, NE speakers maintained duration differences between the members of all contrastive pairs or sets. The two CYG groups produced vowels similarly, and consistently longer in the wordlist-reading compared to the elicitation task. Length differences between the vowels in a contrastive pair were not maintained consistently across tasks or groups. Vowel length differences were maintained in DRESS-NURSE, BATH-TRAP, BATH-STRUT and TRAP-STRUT in the elicitation task, and in KIT-FLEECE and LOT-NORTH in the wordlist-reading task in both groups. CYG trainees also maintained differences between LOT and NORTH in the elicitation task, which was not the case for controls. Furthermore, differences between KIT and FLEECE were also maintained in both tasks when the two groups’ productions were combined.

The comparison with NE speakers’ productions showed that vowel duration in the wordlist-reading task was different between them and CYG learners in most cases, with only KIT, FLEECE, NORTH and GOOSE approximating NE durations. In the elicitation task, vowel duration approximated NE durations in the short vowels DRESS, TRAP, STRUT, LOT and FOOT.

Notably, the overall shorter duration of vowels in the elicitation task seems to have helped learners approximate NE durations in most short vowels, with the exception of KIT. However, these similarities may be due to the task differences between NE and CYG speakers, which were not examined further.

The analysis of CYG vowels showed that, as opposed to English vowels, they were not differentiated by duration, contrary to previous examinations of CYG vowels, which found significant length differences among them (e.g. Themistocleous & Logotheti, 2016; Themistocleous, 2017a, 2017b). This contradictory finding is likely due to phonetic context differences between the current and previous research: whereas CYG speakers in the present study produced their L1 vowels in a /'bVtV/ context, the target vowels in the studies above were flanked by /s/. However, the results of the present study confirm previous findings (e.g. Nicolaidis, 2003; Themistocleous & Logotheti, 2016) that low vowels are longer than high vowels (/a/>/o/>/e/>/u/>/i/). In comparison with the English vowels as produced by the NE speakers, there were significant differences between each CYG vowel and the corresponding English long vowel, whereas no differences were found between CYG vowels and the corresponding English short vowels, which indicates that CYG vowels are more closely associated with short L2 English vowels in terms of duration. It should be noted, however, that due to the lack of a length distinction, CYG vowels may exhibit more durational variability.

Formant analyses confirmed previous findings that due to the small vowel inventory of the language, CYG vowels are well-separated in the acoustic space in stressed positions and in controlled speech (Fourakis et al., 1999; Jongman et al., 1989; Themistocleous & Logotheti, 2016; Themistocleous, 2017a, 2017b). As opposed to this, NE speakers' productions of English vowels were more widely spread in the vowel space, which was not observed in CYG learners' productions of L2 vowels. More specifically, CYG learners produced the L2 vowels of a contrast with considerable overlap, and clustered into the five positions in the vowel space corresponding to the L1 vowels in both tasks, confirming the prediction that CYG learners would use their five L1 vowels for the production of the L2 vowels similarly to the perceptual patterns reported by Georgiou (2019; see 2.5.3). Finally, the productions of each target L2 vowel were very similar in the two tasks and between the two groups, while they were considerably different to NE productions in most cases.

The intelligibility ratings conducted to complement acoustic analyses showed that learners only achieved poor to moderate overall intelligibility, with individual learner ratings ranging from 41%-63% in the elicitation task and 40-71% in the wordlist-reading task. Both CYG groups performed similarly between them and between tasks. The most intelligible vowels produced by CYG learners averaged across the two groups were FLEECE, and the members of the DRESS-NURSE contrast in both tasks. However, above-chance performance was observed in most other vowels as well (TRAP, LOT, FOOT and GOOSE in both tasks, and BATH in the wordlist-reading task). On the other hand, the most problematic vowels were STRUT and KIT, with very low identification scores in both tasks, followed by NORTH in both tasks and BATH in the elicitation task, where identification was below chance levels.

5.1.2 Perception and Production of Contrastive Vowels

Using the assimilation patterns of Georgiou (2019), it was predicted that certain vowel pairs would be perceived and produced as a single L1 category by CYG learners. More specifically, Georgiou (2019) found that L2 English vowels were assimilated to CYG categories as follows: KIT-FLEECE to /i/, DRESS-NURSE to /e/, BATH-TRAP-STRUT to /a/, LOT-NORTH to /o/ and FOOT-GOOSE to /u/. In this section, the perceptual and production performance of participants in each of these contrasts will be examined in detail, in order to provide a comprehensive answer to RQs 1 and 2.

5.1.2.1 KIT-FLEECE

As pointed out by Lengeris (2009a), the English KIT-FLEECE contrast is considered to be the most difficult contrast for L2 learners whose L1 has a small vowel system with no tense-lax or long-short distinctions. Lengeris (2009a) reported that both vowels in the contrast were assimilated to Greek /i/ 100% of the time in his study, and both received high goodness ratings. Georgiou (2019) reported a CG assimilation type for these English vowels on the L1 category /i/, meaning that the two L2 vowels are assimilated to a single L1 category, but with a different goodness of fit. Contrasts belonging to this type of assimilation can be problematic for L2 learners, particularly in the EFL context (Tyler, 2019).

Indeed, the results confirm these expectations, since the KIT-FLEECE contrast was very challenging for these learners, both in production and perception. In production, learners produced the two vowels with considerable overlap in both tasks, merging them in one category closer to FLEECE and CYG /i/, which is higher and fronter than KIT in the vowel space.

This may explain the high intelligibility of FLEECE as opposed to KIT, which cannot be attributed to the duration differences that learners maintained between them, since in that case, we would expect NE raters to identify KIT equally well. The confusion matrix of NE speakers' ratings shows that KIT was mostly identified as FLEECE at a high percentage (69.3% in elicitation and 85% in wordlist-reading), indicating that learners used the same articulatory routines for both vowels. Similar results were obtained in Cebrian (2007), where Catalan learners also had difficulties in the production of the KIT-FLEECE contrast as there was overlap in the acoustic areas of these vowels.

The perception of the two vowels in the contrast was also challenging for learners, as shown by the poor, below-chance identification scores of the vowels. The confusion matrix showed that FLEECE was most commonly identified as KIT followed by the correct vowel, while KIT was most commonly identified as DRESS followed by the correct vowel. Despite the prediction that KIT would mostly be confused with FLEECE, this is not unexpected given the position of KIT between CYG /i/ and /e/ in the vowel space. A similar pattern was found for some Catalan listeners in Cebrian (2006, 2009) and in Fabra (2005) as well as some Danish listeners in Bohn and Steinlen (2003), who also perceived English KIT as similar to their L1 /e/.

5.1.2.2 DRESS-NURSE

According to Georgiou (2019), the two vowels in the DRESS-NURSE contrast may be perceived differently by learners with different proficiency levels. More specifically, Georgiou (2019) found that while these vowels were assimilated to a single L1 category with differing goodness of fit by low proficiency young learners, indicating a CG assimilation pattern, high proficiency young learners showed a different pattern: English DRESS was assimilated to CYG /e/, whereas English NURSE was not consistently assimilated to any CYG category (67% to CYG /e/ and 33% to CYG /a/), leading to the conclusion that in this case, the contrast belongs to a UC assimilation type, whereby discrimination is easier for learners. However, Lengeris (2009a), who examined SMG adults, reported that English DRESS and NURSE were both assimilated to SMG /e/ almost consistently, suggesting a CG assimilation type which can be more challenging for learners. Therefore, predictions about the perceptual and production performance of CYG learners on this contrast were difficult to make, since the learners in the current study were adult intermediate-to-advanced CYG learners of English, unlike both of these previous studies.

In production, the two vowels had the least overlap between them than any other vowel pair, especially in the elicitation task, except in the wordlist-reading productions of the control group. Both vowels were produced close to the CYG vowel /e/, although DRESS had a larger amount of overlap with the CYG vowel in both tasks. The use of CYG /e/ for the production of DRESS was accepted by NE raters as a good instance of the target vowel, and the small modification made for the NURSE vowel seems to have been adequate to lead to high intelligibility ratings by NE speakers for both vowels. This is evident from the intelligibility scores for the two vowels: although DRESS had significantly higher identification scores than NURSE in both tasks, NURSE was also among the most intelligible vowels produced by CYG learners overall. These high intelligibility scores cannot be attributed to duration differences in the productions of CYG learners, since vowel length differences were maintained between them in the elicitation task only, whereas the high intelligibility scores are maintained in the wordlist-reading task as well. Importantly, the differentiation of NURSE from DRESS in the productions of learners seems to be directed towards NE patterns, since there was more overlap between NE speakers and CYG learners' productions in NURSE than in DRESS, particularly in the elicitation task. A similar pattern was observed in the pilot study data (Dimitriou, 2022), where the vowels in the DRESS-NURSE contrast had the largest Euclidean distance between them at T1 compared to other contrasts, with NURSE having the second closest production to NE productions.

As opposed to the KIT-FLEECE contrast, where CYG /i/ was found between the two L2 vowels, CYG /e/ partly overlapped with DRESS and was further away from NURSE, although it was higher than both (lower F1). Therefore, as Georgiou (2019) suggested, it could be predicted that learners would find it easier to perceive the difference between the two vowels. Indeed, DRESS was the third most successfully identified vowel in the perception of CYG participants, although at a moderate rate (50.6%). However, the identification of NURSE was poor (26.5%), eliciting various responses from participants. Contrary to expectations that the two vowels would be confused with each other, the stimuli elicited a variety of responses for these vowels, the second most frequent being STRUT for both. This was consistent with the high proficiency young learners in Georgiou (2019), who classified NURSE as either CYG /e/ or CYG /a/.

Even though assimilation patterns of L2 to L1 categories were not examined in the present study, both perceptual and production results seem to align more with Georgiou's (2019) findings that NURSE is an uncategorised sound. The poor performance of learners in this vowel can be explained by the nature of the task in this study, in which learners were forced to categorise each target vowel using L2 categories. This has likely led to confusion for learners in labelling this sound, as indicated by the variety of responses provided for this vowel. Taking production results into consideration as well, it is hypothesised that if a discrimination test was conducted, participants would have demonstrated a better performance.

5.1.2.3 BATH-TRAP-STRUT

Georgiou (2019) reported an SC assimilation type for these English vowels on the L1 category /a/, meaning that the three vowels are perceived as a single L1 category as equally good or poor exemplars of it. Such contrasts are even more challenging than contrasts belonging to the CG assimilation type, especially for EFL learners (Best & Tyler, 2007; Tyler, 2019), and therefore, it was expected that CYG learners would find these the most difficult vowels to perceive and produce, especially because there are three members in the contrast.

Indeed, even though there was some overlap between BATH-STRUT and TRAP-STRUT in the productions of NE speakers as well, the three vowels and CYG /a/ overlapped almost completely in the production of learners in both tasks, but especially in the wordlist-reading task. Furthermore, while length differences were maintained between the three vowels in the elicitation task, they had similar durations in the wordlist-reading task, where all were significantly different to NE productions. Overall, learners' productions in the elicitation task were closer to NE speakers' productions in all three vowels, both in spectral characteristics and duration; however, this was not reflected in the intelligibility scores for the three vowels, where the task effect was significant only in BATH, in which case participants' productions in the elicitation task were perceived less accurately by NE raters.

In addition, CYG learners' productions of BATH and STRUT had more overlap with NE productions compared to TRAP. It should be noted, however, that whereas BATH took up a very limited area of the vowel space in NE speakers' productions, CYG learners' productions were more widely spread, possibly explaining the high degree of overlap in this case. This is consistent with Baese-Berk et al.'s (2020) observation that non-native speakers may exhibit

less “articulatory stability”, i.e. more variable productions compared to native speakers, due to the fact that they have less experience in performing the articulatory gestures for L2 sounds. For instance, Wade et al. (2007) found that Spanish learners of English had more variable productions in L2 English vowels compared to NE speakers.

This wide variability found in the production of BATH by CYG learners, which contributed to the overlap with NE speakers’ productions, led to overlap with other vowels as well (i.e. TRAP, STRUT and LOT), which possibly affected the identification performance of NE raters in this vowel. Indeed, the confusion matrix shows that BATH was mostly perceived as LOT by NE raters in the elicitation task, where the vowel was identified at a low rate (29%). The above-chance identification of BATH in the wordlist-reading task, however, was unexpected. A closer examination of the ratings for this vowel showed that the higher identification scores were mostly due to trainees’ productions in this task, even though wordlist-reading productions of BATH had more overlap with both TRAP and STRUT, as well as less overlap with NE speakers’ productions than in the elicitation task. Since vowel length differences were not maintained between the vowels in this contrast either in this task, the only possible explanation could be the fact that despite efforts to remove the post-vocalic /r/ in the stimuli for BATH (i.e. “Bart” and “gart”), trainees may have produced the approximant in such a way as to affect the quality of the preceding vowel. However, this explanation cannot be confirmed, since no analyses on the production of post-vocalic segments was conducted in either group’s productions, and there is nothing to suggest that controls’ productions did not involve the same features.

The high degree of overlap between NE and CYG speakers observed in STRUT can be explained by a comparison between L1 and L2 vowels: more specifically, CYG /a/ had more overlap with STRUT than the other vowels in the contrast, and therefore, the use of the L1 vowel in learners’ productions corresponded to this L2 vowel more than the other two. However, despite the overlap with NE speakers’ productions, this vowel had the poorest identification of all vowels overall by NE raters, who mostly perceived it as TRAP or LOT. Finally, intelligibility scores showed that TRAP, which had less overlap with NE productions, was identified significantly better by NE raters, indicating that this was perceived as a better exemplar of the L2 vowel.

In addition to difficulties in production, learners evidently struggled to perceive the three vowels as well, with poor identification scores in BATH (18.8%) and STRUT (32.4%), and

moderate, above-chance identification for TRAP (43.1%). The confusion matrix showed that TRAP and STRUT were mostly confused with each other, with BATH as a third option with lower frequency. BATH was also confused with STRUT and TRAP, but it was perceived as LOT in some cases as well. This is consistent with the assimilation patterns observed in Lengeris (2009a) for SMG adult learners but not Georgiou (2019) for CYG young learners, where SMG learners assimilated English TRAP and STRUT to the same L1 vowel, i.e. SMG /a/, and English BATH, LOT and NORTH as SMG /o/.

Overall, learners found the perception and production of these vowels challenging, as predicted, and this was also reflected in the ratings of NE raters. Of the three vowels, TRAP was better perceived by learners and better identified by NE raters, possibly due to the similarity between TRAP and /a/ in terms of vowel height (higher F1 values for both, indicating lower vowel). These results suggest that the three vowels are merged into a single category corresponding to L1 /a/, which is mostly perceived by NE speakers as TRAP or LOT, presumably depending on vowel backness (i.e. more fronted productions are perceived as TRAP, and backer productions are perceived as LOT). The difficulties observed in this contrast were expected, given that the L2 contains three vowels that can be perceived as a single L1 category, making it even more difficult to detect the subtle differences between them.

5.1.2.4 LOT-NORTH

Similarly to the BATH-TRAP-STRUT contrast, Georgiou (2019) reported an SC assimilation type for LOT-NORTH as well, whereby the two L2 vowels are assimilated to the L1 category /o/. Therefore, it was predicted that learners would struggle to perceive and produce the vowels in this contrast as well.

As in other vowel contrasts, learners used their L1 vowel for the production of both members of the LOT-NORTH contrast as well, merging them in a single category despite the fact that one member is closer to the L1 category than the other. This also led to a higher degree of overlap between CYG learners and NE speakers' productions of LOT than NORTH, especially in the elicitation task, as well as significantly higher intelligibility ratings for LOT than NORTH in both tasks. Confusion patterns showed that while LOT was identified correctly most of the time, NORTH was frequently mis-identified as LOT by NE speakers, suggesting that while the use of CYG /o/ is an acceptable production for LOT, it is not for NORTH. Duration differences between the two vowels in the productions of CYG learners were maintained in

the wordlist-reading task for both groups, but only for trainees in the elicitation task. It should be noted that despite NE speakers' reliance on spectral cues to distinguish the vowels of a contrast, with duration being a secondary cue (e.g. Hillenbrand et al., 2000), both acoustic cues might be used to differentiate the members of a contrast (Zhang et al., 2021). In this case, the length between the vowels may have facilitated the identification of NORTH in wordlist-reading productions, which were found to be more intelligible in this task than in the elicitation task.

In perception, LOT was the fourth most successfully identified vowel in the perception of CYG learners, although at a moderate, below-chance rate (47.1%), while the identification of NORTH was poor (35.4%). Despite the fact that CYG /o/ is very close to English LOT and much lower than English NORTH, participants found it challenging to differentiate between the two in perception, as shown by the confusion matrix where the second most frequent response was NORTH (30%) and LOT (21.1%), respectively, although NORTH was also occasionally confused with the back vowels FOOT (16.6%) and GOOSE (19.7%), which is understandable given the overlap of English NORTH to CYG /u/.

5.1.2.5 FOOT-GOOSE

Finally, as per Georgiou's (2019) assimilation patterns, the vowels in the FOOT-GOOSE contrast are assimilated to CYG /u/ with different goodness-of-fit ratings, therefore forming a CG assimilation type. However, it should be noted that spectrally, the two vowels were fronted in the productions of NE speakers in the present study, a common feature in contemporary SSBE, which renders the two vowels very different compared to the CYG high back vowel /u/.

The productions of CYG learners for this contrast show that the two vowels almost completely overlapped, particularly in the elicitation task, since both members of this contrast were produced using learners' L1 /u/, which is spectrally very different to the L2 vowels, particularly in terms of backness (lower F2 values). Evidently, the two vowels were perceived by learners as sufficiently similar to their L1 category, leading to the use of the L1 articulatory gestures for the production of the L2 vowels. Inevitably, the productions of CYG learners for both vowels were very different compared to NE speakers, possibly due to a lack of input received by CYG learners from this variety. Despite the considerable differences between CYG and NE productions, the two vowels were identified correctly at an above-chance level by NE raters, although at a moderate rate. Confusion matrices showed that the second most

frequent response of NE raters was the other vowel of the contrast, suggesting that there was confusion between them in identification, i.e. raters were able to perceive that the productions were of a high back vowel, but could not identify which of the two was intended. Despite the spectral similarity between CYG /u/ and NORTH, NE raters did not identify either of the back vowels as NORTH to a high percentage. At the same time, duration analyses show that the learners produced the two vowels with similar durations, indicating that duration was not a cue that NE raters could have used in identifying the intended vowel.

Despite using the L1 vowel in their productions, the vowels of this contrast had the highest identification scores in the perception of CYG learners, although at a moderate rate (57.9% for FOOT and 59.2% for GOOSE). This was the only contrast where learners were able to identify both members at an above-chance level, with the second most frequent response being the other member of the contrast, although at a very low percentage. The higher performance of learners in perceiving the vowels of this contrast may be due to the fact that they are both spectrally very different to any CYG vowel; learners may therefore be able to perceive the subtle differences between them in perception, despite their inability to produce them accurately. Another possibility is that the presence of the post-vocalic /l/ in one of the stimuli used for GOOSE (i.e. “gould”) and its effect on the quality of the preceding vowel may have facilitated learners in the identification of the target vowel in this pair. This limitation will be further discussed in 6.1.

5.1.3 Summary of Findings

Overall, both vowels in the pairs KIT-FLEECE and LOT-NORTH were challenging in perception, while the production of one of the vowels in each pair (i.e. FLEECE and LOT) as rated by NE naïve listeners was significantly higher than the other, since the L1 vowel used by CYG learners in production was closer to the respective L2 vowel. A similar pattern was observed in BATH-TRAP-STRUT, where TRAP was better perceived and produced by learners, despite the higher spectral overlap between the L1 vowel and STRUT.

In the DRESS-NURSE contrast, even though DRESS was better perceived by learners and better identified by NE raters than NURSE, the latter also had one of the highest intelligibility scores. Despite the fact that learners used their L1 vowel /e/ in the production of DRESS and they only slightly differentiated their productions for NURSE, both vowels were accepted as good instances of the target L2 vowels by NE raters. As the only L2 English vowel that is not

consistently assimilated to a single L1 vowel, and based on the results of this study as well as Dimitriou (2022), it can be concluded that CYG learners are likely to form a new category for NURSE, provided that sufficient input is received that can help them better attune their perception and production to the specific acoustic cues of this vowel.

Finally, FOOT and GOOSE were almost equally distant from CYG /u/, which means that despite the fact that learners may be able to discern the differences between the L1 and L2 vowels, they seem to be unable to distinguish between the two L2 vowels, and therefore use a single vowel corresponding to the L1 /u/ vowel in the production of both of them, eliciting moderate intelligibility scores for both by NE raters.

5.1.4 General Discussion

The findings reported above are in agreement with current models of speech production and perception (SLM, SLM-r, PAM and PAM-L2) and previous findings arguing that the relationship between the L1 and L2 sound inventories is important in identifying the difficulties faced by L2 learners (e.g. Aliaga-Garcia & Mora, 2009; Alispahic et al., 2017; Lengeris & Hazan, 2007, 2010; Lengeris, 2018, among others), as well as the argument that learners will find it more difficult to avoid using an L1 instead of an L2 sound if there is a close association between them, due to the fact that replacing old habits is more difficult than learning a new set of language habits (Koutsoudas & Koutsoudas, 1962). CYG learners perceived L2 English vowels poorly to moderately, suggesting that they cannot easily differentiate the members of an L2 contrast, and therefore, new category formation is blocked due to equivalence classification.

Learners' difficulties in perceiving the subtle differences between members of a contrastive pair were expected, given that most assimilation types reported in Georgiou (2019) belong to the CG or the SC type. According to the PAM-L2, vowels that belong to these assimilation types are challenging for L2 learners, but as Tyler (2019) points out, they are even less likely to be acquired in an EFL context, especially if the foreign-accented input received does not differentiate the L2 phonemes, as is likely in the case of CYG learners (see 5.4.2).

Furthermore, CYG learners produced the L2 vowels using the same articulatory routines as for L1 sounds, which were rated as good instances for some vowels but not others. The use of unmodified L1 vowels in English words can be accepted by NE listeners as good exemplars of the target vowel and may go unnoticed (Flege & Wayland, 2019), as has also been observed

in other studies, such as Cebrian (2007), where NE listeners perceived the Catalan vowels /i/, /ɛ/ and /ei/ as good instances of the acoustically closest English categories /i/, /ɛ/ and /eɪ/.

As concerns the perception-production link, pre-test results do not allow for any robust conclusions to be drawn: whereas in some vowels, low perceptual performance was associated with low intelligibility scores (e.g. KIT, BATH, STRUT, NORTH) at least in one of the two production tasks, in others, low or average perceptual performance was associated with higher intelligibility scores (e.g. FLEECE, DRESS, NURSE, TRAP, LOT). The highest, albeit still moderate, perceptual score was achieved for FOOT and GOOSE, which was associated with moderate intelligibility scores as well. However, spectral analysis demonstrated that these vowels are among the least native-like, suggesting that new category formation has not taken place for these vowels either.

Furthermore, individual participant performance in perception and production were not necessarily aligned as opposed to the findings of Melnik-Leroy et al. (2022), who found that learners who were good producers were also good perceivers, while only half of the good perceivers were also good producers. Based on this finding, the researchers argued that "good production very rarely occurs in the absence of good perception while good perception often occurs in the absence of good production" (p. 598). In the present study, however, this was not the case, since correlation analysis in participants' individual performance between perceptual and production tasks showed no significant correlation between them.

Based on the assumptions of the SLM, it could be argued that learners' overall insufficient perceptual ability hindered their production performance as well, allowing them to produce intelligible vowels only when the L1 vowel was judged to be a good exemplar of the corresponding L2 vowel. Based on this assumption, the fact that learners did not discern the subtle phonetic differences between similar L1 and L2 sounds prevented them from forming new phonetic categories that could lead to a more accurate production of L2 vowels. However, the case of NURSE is a notable exception; while learners demonstrated poor perceptual identification of the vowel, they were able to modify their productions towards native-like patterns that were judged as highly intelligible by NE raters. At the same time, it would be a stretch to argue that accurate production was achieved despite inaccurate perception (as in Bohn & Flege, 1997) for the vowels that reached high intelligibility scores (FLEECE, DRESS, NURSE, TRAP and LOT), because spectral analysis showed that learners

merely produced these vowels using their L1 articulatory routines, which were simply perceived as better exemplars of the corresponding L2 vowels, while at the same time, they remained different to NE speakers' productions.

Therefore, the results of this study are more aligned with previous studies that reported mixed results, such as better perception for some sounds and better production for others (e.g. Hao & de Jong, 2016) or which report a lack of an overall significant correlation between perception and production data (e.g. Fabra & Romero, 2012; Kartushina & Frauenfelder, 2014; Peperkamp & Bouchon, 2011). Based on this finding, the results of the current study seem to fit better to the assumptions of the revised model (SLM-r) that "production and perception coevolve without precedence" (Flege & Bohn, 2021, p. 64).

Finally, the presence or absence of orthographic cues did not seem to have affected participants' productions to a large extent overall. More specifically, spectral analysis showed a large degree of overlap between the vowels produced in each task. Spectral analysis indicated that more vowels were closer to NE productions in the elicitation task, suggesting that participants may have benefited from listening to and imitating NE speakers in their productions rather than reading the orthographic forms, similarly to participants in Detey et al. (2014) who were also found to perform better in the repetition task rather than the reading task. However, the effect of task on intelligibility data only reached significance in three vowels, two of which were rated as more intelligible in the wordlist-reading task and one as more intelligible in the elicitation task. Furthermore, there was no effect of task on the intelligibility scores of individual participants or of group performance at T1. The lack of an effect of task was unsurprising given the nature of elicited imitation tasks, which confound perception and production skills, meaning that learners' productions still depended on how accurately the stimulus was perceived (Nagle & Baese-Berk, 2022). Since perception was moderate, any advantage potentially offered by an elicitation task was lost.

The effect of task was mostly evident in the duration of vowels, most of which were significantly longer in the wordlist-reading task. However, this was expected due to the nature of the task, and did not affect intelligibility ratings, since NE raters did not appear to rely extensively on duration as a cue to distinguish between the members of any contrastive pairs, except in LOT-NORTH, where duration differences between the two vowels may have facilitated the identification of the vowels in wordlist-reading productions.

Having discussed the findings in relation to RQs 1 and 2 which relate to data obtained at T1, this discussion now turns to the effects of HVPT observed in this study, and therefore focuses on T2 data and their comparison with T1 data. This second part of the discussion will address RQ 3 "How effective is HVPT in improving CYG learners' vowel perception and production without explicit production training?". It begins with a summary and discussion of the findings in relation to learners' perceptual and production performance at the two time points, including the results of the perceptual identification task, spectral and durational analyses, and intelligibility ratings at a group and individual learner level as well as performance in individual target vowels. Then, a general discussion of the findings in relation to the RQ and previous findings is provided.

5.2 Effects of HVPT on the Perception and Production of English Vowels

5.2.1 Perception

Perceptual results showed that both groups' overall perceptual performance improved from T1 to T2, without significant differences between them at either time, similarly to the findings of Cebrian et al. (2019). Therefore, trainees' better perceptual performance at T2 cannot be attributed to the training alone. This improvement in both groups from T1 to T2 may be the result of additional input received from other sources between the two sessions, during which participants may have paid more conscious effort to the perception of L2 vowels due to an interest developed through their participation in the study. However, it should be noted that in the examination of individual participant performance, it was observed that six of the eight trainees improved significantly in their perceptual performance from T1 to T2, while only one of the six controls performed significantly better at T2. Similar results are reported in Cebrian et al. (2019), who found that the performance of the control group in a perceptual identification task was also improved at T2, although at a smaller rate. Overall perceptual performance at T2 ranged from poor for some participants (e.g. CYG10, CYG11, CYG16) to good for others (e.g. CYG04, CYG08, CYG12), ranging from 10.1% to 72.3%.

Differences between the two groups at T2 were mostly observed in their perception of individual vowels. More specifically, while there were no differences between the two groups in the identification of each target vowel at T1, trainees perceived the vowels FLEECE and NURSE significantly better than controls at T2. Furthermore, trainees significantly improved in the identification of multiple vowels (KIT, FLEECE, BATH, TRAP and NORTH) across time,

whereas controls only improved significantly in one vowel (KIT), although the perception of these vowels remained moderate for both groups (range: 43.8%-63%).

Even though some studies (e.g. Wang & Munro, 2004) report that identification training with feedback can be effective in improving all contrasts included in the training, this improvement in some vowels but not others is consistent with other studies that report partial improvements as an effect of training. For example, Georgiou (2021) reported that the HVPT administered to learners was effective in improving trainees' identification accuracy, but he observed that some vowels, i.e. TRAP, BATH and LOT, improved more than others, even though the study does not specify whether these improvements were statistically significant. A similar finding is reported in Carlet and Cebrian (2014), where participants significantly improved in the identification of some of the trained segments but not in others.

It should be noted that the improvements observed cannot be attributed to familiarity effects, since some participants had a lower overall perceptual performance at T2, while the identification of some vowels was also lower at T2 compared to T1.

5.2.2 Production

5.2.2.1 Duration

As with T1 data, vowels in the wordlist-reading task were consistently longer than vowels in the elicitation task. The two groups had no significant differences between them in their productions in each task. However, they did have differences in whether they maintained duration differences between the members of a contrast. Irrespective of task, trainees maintained differences in DRESS-NURSE, BATH-TRAP, BATH-STRUT and LOT-NORTH; in all cases the long vowel was longer than the short vowel at T1 and the difference was maintained or increased at T2. Controls, on the other hand, only maintained differences between DRESS-NURSE, where the long vowel was significantly longer than the short vowel in the pair.

Duration analyses in relation to this RQ mostly focused on three comparisons, i.e. the productions of trainees compared to controls, the productions at T1 compared to T2, and the productions of NE compared to CYG speakers. In the comparison between trainees and controls across time, the two groups had differences in the duration of STRUT in the elicitation task and in KIT, DRESS and LOT in the wordlist-reading task, where trainees shortened these short vowels, while controls lengthened them at T2.

In the comparison between T1 and T2 productions of learners, it was observed that trainees did not alter their vowel durations significantly at T2 in the elicitation task, but they did so in the wordlist-reading task in DRESS-NURSE and LOT-NORTH, where they produced the vowels in these pairs with an increased distance between the short and long vowel. Controls, on the other hand, increased the duration of STRUT in the elicitation task, decreasing its distance from BATH, which was shortened, contrary to NE patterns, while they did not alter their vowel durations at T2 compared to T1 in the wordlist-reading task.

Finally, the comparison between trainees and NE speakers showed significant differences in NURSE, BATH, STRUT and LOT, since NE speakers maintained more distance between short and long vowels, i.e. they produced the long vowels with significantly longer durations and the short vowels with significantly shorter durations than trainees in the wordlist-reading task. Controls had significant differences with NE speakers in most vowels in the wordlist-reading task (DRESS, NURSE, BATH, TRAP, STRUT, LOT and NORTH). Despite task differences between the elicitation task completed by CYG learners and the wordlist-reading task completed by NE speakers, the comparison between participants' productions showed that both CYG groups had significant differences with NE speakers in the same, long vowels (NURSE, BATH, NORTH and GOOSE), meaning that they approximated NE durations in the same vowels as at T1, with the addition of KIT and FLEECE.

5.2.2.2 Spectral Characteristics

In terms of spectral characteristics, it was evident that learners' productions remained almost the same at T2, with Pillai scores close to 0 in most cases, meaning that the training was insufficient to alter learners' use of spectral cues for the production of L2 vowels. Some noteworthy differences between T1 and T2 include the production of FOOT by the control group in the wordlist-reading task, which was also associated with significant changes in intelligibility, as will be further discussed in 5.2.3.3, and of the DRESS-NURSE contrast by both groups. More specifically, trainees differentiated the two vowels in the contrast in both tasks and times, while controls merged the two vowels at T2, even though they were differentiated in the elicitation task at T1. However, these modifications did not lead to significant changes in the intelligibility scores of the two vowels across time.

5.2.3 Intelligibility Ratings

5.2.3.1 Global Intelligibility

In terms of intelligibility, it should be noted that while there were no differences between the two tasks at T1, both trainees and controls received higher identification scores in the wordlist-reading task than in the elicitation task at T2. This is contrary to Detey et al. (2014), where participants were found to produce the target words better in the repetition than in the reading task. However, it can be explained by the fact that the elicited imitation task is more challenging than reading tasks, as noted by Thomson and Derwing (2016).

Furthermore, while the two groups of learners performed similarly at T1, trainees significantly outperformed controls at T2, especially in the wordlist-reading task. However, the training was not adequate to cause any significant improvements in the overall performance of either trainees or controls between T1 and T2; on the contrary, controls' performance in the elicitation task was significantly lower at T2 compared to T1. Overall intelligibility was moderate, with the lowest percentage observed in controls in the elicitation task at T2 (46%), and the highest percentage reaching only 58% in trainees in the wordlist-reading task at T2.

5.2.3.2 Individual Participant Performance

When participants were examined individually, it was established that only two participants (CYG05 and CYG15; one trainee and one control) performed differently at T2 compared to T1 in the elicitation task, and only one (CYG16; trainee) in the wordlist-reading task. However, only the trainee in the wordlist-reading task performed better at T2 compared to T1, even though his performance remained moderate (52%). The two participants who had a different performance in the elicitation task received lower intelligibility scores at T2, despite the fact that one of them was a trainee. Taking into consideration the fact that other participants, including both trainees and controls, received lower ratings at T2 compared to T1 as well, albeit insignificant in other cases, it is hypothesised that participants may have been overwhelmed or confused upon realising the existence of such vowel contrasts in English. In combination with a lack of adequate, native-like input, this may have resulted in an unsuccessful attempt to identify and use the appropriate cues to differentiate between the members of the contrast in their productions, whether they had received training or not.

Indeed, Flege and Wayland (2019) point out that establishing a new phonetic category involves specifying how multiple cues are integrated and weighted, which can be challenging

for learners when the weighting of these acoustic cues is different in the L1. Furthermore, as Flege and Bohn (2021) note, the time course of L2 category formation is not well understood yet, which makes it possible that upon identification of differences between L1 and L2 segments or between contrastive L2 sounds, learners may shift their productions in an experimental way, before reaching more accurate productions at a later stage with additional input.

5.2.3.3 Intelligibility in Individual Vowels

Despite the large overlap between T1 and T2 productions, intelligibility ratings suggest that some vowels were perceived differently at the two time points. In the analysis of the intelligibility of target vowels by learners, only a few vowels were found to be improved at T2. In trainees, only BATH was improved in the elicitation task, and only FLEECE in the wordlist-reading task. At the same time, LOT and FLEECE in the elicitation task became significantly less intelligible. The performance of controls was significantly improved in the production of FOOT in the wordlist-reading task between T1 and T2.

The reasons behind these differences are unclear, given the lack of any substantial changes in the productions of learners that could lead to a change in the identification of these vowels. To begin with, there was a high degree of overlap between the T1 and T2 productions of the BATH vowel by trainees in the elicitation task (Pillai score 0.006), while the plot in **Figure 19** shows that the vowel was produced very similarly at the two time points. Furthermore, there was no significant effect of time on the duration of this vowel or the other vowels in the contrast (TRAP, STRUT) that could provide a secondary cue to NE raters; length differences between the members of the contrast were maintained at both T1 and T2. Therefore, it is hypothesised that the higher intelligibility of this vowel was due to the fact that there was slightly less variability in the production of BATH at T2, in combination with the fact that it was further separated from TRAP, STRUT and LOT at T2. Indeed, confusion matrices (see **Table D1** and **Table D5** in Appendix D) on the ratings of NE speakers demonstrate that no instances of BATH at T2 were perceived as STRUT (as opposed to 10% at T1), while the perception of BATH as LOT or as TRAP was also reduced (40% at T1, 23.1% at T2 for LOT; 16% at T1, 12.3% at T2 for TRAP).

As concerns the improvement in FLEECE by trainees in the wordlist-reading task, spectral characteristics cannot explain why NE raters perceived the vowel better at T2, since there was

a high degree of overlap between T1 and T2 productions (Pillai score 0.0283), and the vowel was almost completely merged with KIT at both times (see **Figure 19**). Furthermore, there were no duration differences between T1 and T2 productions. The only difference that could account for the higher intelligibility of this vowel is the fact that KIT was significantly shortened at T2, which may have led to fewer instances of FLEECE being misidentified as KIT, as confirmed through the confusion matrices (**Table D3** and **Table D7** in Appendix D) that show that FLEECE was misidentified as KIT at a lower percentage at T2 (14.7% at T1 compared to 10% at T2). However, there was no improvement in the identification of KIT, and therefore, the reasons behind the higher intelligibility scores for FLEECE remain unclear. It should be noted, however, that the identification of this vowel was very good at both times (80% at T1 and 89% at T2).

The reasons for the improvement in the intelligibility of FOOT in controls' productions in the wordlist-reading task are also unclear. More specifically, no differences were found in the duration of the vowel between T1 and T2, and duration differences with GOOSE were not maintained at either time. As concerns spectral characteristics, the T1 and T2 productions of FOOT by the control group in the wordlist-reading task had the least overlap between them (**Table 17**), showing that T2 productions were different to T1 productions. The plot in **Figure 19** also shows that both FOOT and GOOSE were backer at T2. These small differences in spectral characteristics may have contributed to the higher intelligibility score of this vowel at T2, which was at a very good rate (83%) despite being less native-like. The reasons behind controls' modification of their productions, however, remain unclear; based on Flege and Bohn's (2021) argument (see 5.2.3.2) it can be hypothesised that this modification may be the result of learners attempting to alter their productions in an experimental way upon becoming aware of differences in the vowels of the L2.

The vowels that received lower intelligibility scores at T2 compared to T1, i.e. LOT and FLEECE, both belong to the productions of trainees in the elicitation task. These were also very similar in spectral characteristics at the two time points, as were their counterparts, NORTH and KIT, respectively. However, it can be seen from **Figure 19** that both of these vowels become more merged with the second vowel in the pair at T2 compared to T1. Even though this is a small difference, it may have contributed to the difficulties of NE raters in identifying the intended vowel. Indeed, confusion matrices (**Table D1** and **Table D5** in Appendix D) show that both LOT and FLEECE were misidentified as NORTH and KIT, respectively, more frequently

at T2 compared to T1: LOT was misheard as NORTH 14.7% of the time at T1 compared to 24.3% at T2; FLEECE was perceived as KIT at a percentage of 13.3% at T1 compared to 24.3% at T2. Duration did not seem to play a role in the perceptions of NE raters, since there were no significant differences in the duration of these vowels or their counterparts between T1 and T2.

Finally, despite the fact that elicitation tasks involve an element of perception (Nagle & Baese-Berk, 2022), participants' perception of these target vowels did not seem to have played a role in their productions in this task. More specifically, while trainees improved in the perception BATH at T2, which may have caused the small modification of their productions leading to higher intelligibility scores in the elicitation task, the same was not observed for FLEECE, where perceptual improvements did not lead to modifications of this vowel in production, but rather were associated with lower intelligibility scores. Furthermore, no changes occurred in trainees' perception of LOT, while intelligibility ratings were lower.

5.2.4 Summary of Findings

In summary, the findings suggest that the training was not adequate to lead to significant improvements in trainees' overall perceptual or production performance. However, similarly to other studies (e.g. Carlet & Cebrian, 2014; Georgiou, 2021), some improvements in perception or production were observed in some trained segments but not others. In perception, trainees improved in more vowels than controls (KIT, FLEECE, BATH, TRAP and NORTH as opposed to only KIT for controls) and identified two vowels (FLEECE and NURSE) better than controls at T2, although the perception of these vowels remained moderate even after improvement.

Acoustic analyses showed that trainees had more native-like tendencies in the duration of vowels than controls, as they increased the distance between the short and long vowels in some pairs, while controls reduced that distance at T2. Furthermore, while trainees' vowel duration was different to NE speakers in four vowels (NURSE, BATH, STRUT and LOT), controls had differences with NE speakers in most vowels (DRESS, NURSE, BATH, TRAP, STRUT, LOT and NORTH). Spectral characteristics remained very similar between T1 and T2, meaning that the training was insufficient to alter learners' use of spectral cues for the production of the target vowels. Some modifications in the productions of learners in FOOT-GOOSE and DRESS-NURSE

are noted, but these did not affect intelligibility ratings, with the exception of FOOT by the control group in the wordlist-reading task.

Intelligibility ratings showed mixed results for individual vowels, with better ratings for some vowels and worse for others. FLEECE, DRESS and NURSE remained high in intelligibility, with the exception of NURSE in the wordlist-reading task of controls, who merged DRESS and NURSE in their productions at T2, while KIT and STRUT remained the most problematic vowels across group, task and time. Trainees improved in their intelligibility of BATH in the elicitation task and FLEECE in the wordlist-reading task, but produced LOT and FLEECE in the elicitation task with lower intelligibility at T2. Controls showed improved intelligibility in their production of FOOT in the wordlist-reading task.

In addition to varying results in different vowels, the performance of individual participants also varied in both perceptual performance and intelligibility ratings. In perception, five of the eight trainees (CYG03, CYG04, CYG08, CYG12 and CYG17) and one of the six controls (CYG15) improved, and one trainee (CYG16) performed worse at T2. In intelligibility, the trainee who performed worse in perception (CYG16) had a higher intelligibility rating in the wordlist-reading task, while the control that improved in perception (CYG15) and a trainee whose perceptual performance was not changed (CYG05) had lower intelligibility scores in the elicitation task. The perceptual and intelligibility scores of all three participants (CYG05, CYG15 and CYG16), however, remained moderate to poor across time. These mixed results obtained for individual participants lead to the conclusion that individual differences may have a role to play in their performance; these will be addressed in 5.4.

5.2.5 General discussion

In their meta-analysis, Sakai and Moorman (2018) showed that perception-only training can yield medium-sized gains for perception and at least small production gains, suggesting that it is possible for perception training alone to yield a small but robust improvement in the production modality. This was not observed in the overall perceptual or production performance of learners in the present research, even though perceptual gains can be reported for most trainees. As opposed to previous studies showing that the perceptual gains from HVPT can be transferred to production (e.g. Bradlow et al., 1997; Huensch & Tremblay, 2015; Jügler et al., 2015; Lambacher et al., 2005; Lengeris & Hazan, 2010; Motohashi-Siago & Hardison, 2009; Okuno & Hardison, 2016; Rato & Rauber, 2015; Shinohara & Iverson, 2015;

Thomson, 2011; Wong, 2013, 2015), this study found only partial improvement in very few vowels, which was mostly evident in intelligibility ratings, and which was so sporadic as to suggest that it may have been due to factors other than the training. This is in line with various previous studies (e.g. Aliaga-Garcia & Mora, 2009; Garcia Perez, 2003; Iverson et al., 2012; Lopez-Soto & Kewley-Port, 2009; Peperkamp & Bouchon, 2011; Zhang et al., 2021) which found little or no improvement in production at all. Therefore, this study cannot provide support to the argument that perception training can automatically lead to production improvement.

However, as in Aliaga-Garcia and Mora (2009), even though no overall significant gains in perceptual or productive performance were observed for all vowels examined, there was significant improvement in either perceiving or producing some of the target sounds after training, suggesting that phonetic training may have different effect sizes on learners' perceptual and productive competence based on phonetic dimension and sound contrast. Therefore, the results of this study confirm Wade et al.'s (2007) observation that HVPT may yield different benefits for different vowel categories, or even no benefits at all for highly confusable vowels.

The finding that learners' productions remained almost the same at T2 is contrary to studies involving SMG learners (Lengeris & Hazan, 2010; Lengeris, 2009a, 2018), which found that while learners used their five L1 categories in English vowel production before training, the overlap between the L2 vowels was much less after the training, suggesting that following perceptual training SMG learners learnt to differentiate English vowels in their production as well. Moreover, these studies noted much less overlap in FLEECE-KIT, DRESS-NURSE and BATH-TRAP-STRUT at T2, which was not found in the present research. However, the results of the present study are in line with previous results involving CYG learners, and more specifically Georgiou (2021) who trained both children and adults and found that while HVPT was effective in improving trainees' identification accuracy, transfer of perceptual gains to production was only significant in children and not adults, according to NE raters' judgements.

Furthermore, Lengeris (2009a) and Lengeris and Hazan (2010) noted that L2 vowel perception and production were aligned, but only after participants were exposed to large amounts of variable L2 input, a finding that supported the link between perception and production. Similarly, many studies reported at least a modest or partial relationship between

the two modalities (e.g. Baker & Trofimovich, 2006; Bettoni-Techio et al., 2007; Jia et al., 2006; Kluge et al., 2007; Levy & Law, 2010; Melnik-Leroy et al., 2022; Zhang & Peng, 2017). Rather, the results of the present research are in agreement with studies that found no correlation between the two modalities (e.g. Huensch & Tremblay, 2015; Kartushina & Frauenfelder, 2014; Peperkamp & Bouchon, 2011), since perceptual performance did not seem to be aligned with production performance on any level, including overall performance, individual vowel performance or individual participant performance.

These findings are therefore more in line with Flege and Bohn's (2021) revised hypothesis that the two modalities may be developed independently from one another. Although it could be hypothesised that learners simply did not update their productions even when accurate perception of a sound was achieved, as argued in the initial formulation of the SLM, the findings show that production performance was more successful than perceptual performance in some learners and target vowels, indicating that perceptual learning does not necessarily precede production learning. However, it is crucial to remember that evaluating perception and production performance, and the link between them is complicated, and it involves numerous methodological decisions that may affect the outcomes, especially when perception and production data are compared, in which case tasks may not involve equally demanding tasks for the two modalities (Melnik-Leroy et al., 2022; Nagle & Baese-Berk, 2022).

While the results of this study may seem discouraging, they were not surprising. Similarly to learners in previous studies (e.g. Aliaga-Garcia & Mora, 2009), CYG learners have been exposed to the L2 mostly through formal classroom instruction in an EFL context, where L2 input is limited and usually foreign-accented (see 2.2 and 5.4.2), which makes it unlikely for learners to develop L2 phonetic categories for the target sounds, as evident by their T1 performance as well. It is also of crucial importance to note that HVPT studies vary widely in terms of methodological choices and training paradigms used, which can play a key role on the outcomes on perceptual and production performance and can explain, at least partly, the inconsistency in the results found in previous studies, as pointed out by many researchers (e.g. Barriuso & Hayes-Harb, 2018; Hu et al., 2022; Melnik-Leroy et al., 2022; Nagle & Baese-Berk, 2022; Sakai & Moorman, 2018; Thomson, 2018).

The lack of an effect of the training on overall production and perception performance in this study is likely due to methodological choices, including the length of the training in

relation to number of target segments, the training and testing methods used or the types of analyses conducted (Nagle & Baese-Berk, 2022). Differences in the results can also be attributed to individual differences among participants both within the present study and across studies, as well as their L2 experience or L1 background, as will be discussed in 5.4 (Barriuso & Hayes-Harb, 2018; Kartushina & Martin, 2019; Nagle & Baese-Berk, 2022).

One important aspect of the training that may have hindered learning, i.e. the high-variability stimuli provided, merits more attention here. More specifically, despite arguments in favour of high-variability stimuli in promoting the generalisation of learning (e.g. Brosseau-Lapré et al., 2013; Qian et al., 2018; Thomson & Derwing, 2016; Thomson, 2011, among others), variability was found to impede learning when the target L2 contrasts are difficult or highly confusable in relation to the learners' L1 (e.g. Giannakopoulou et al., 2017; Wade et al., 2007) or when learners are weak or novice learners, in which case the added processing costs required to process speech by multiple talkers may prove detrimental (e.g. Antoniou & Wong, 2015; Chang & Bowles, 2015; Perrachione et al., 2011; Sadakata & McQueen, 2014).

Furthermore, Kartushina and Martin (2019) and Evans and Martín-Alvarez (2016) point out that variability in perceptual training may negatively affect production, as the researchers found higher production improvement in a single-talker than in a multiple-talker condition. This suggests that LVPT may be more beneficial for some learners when the goal is promoting production improvement. The inclusion of high variability in both testing and training the learners in this study is very likely to have negatively affected the perceptual and production performance of at least some participants. This is evident by the examination of participants' individual performance; more specifically, some participants' performance was lower at T2 compared to T1 in both perception and production. It is therefore hypothesised that these learners may have benefitted more from training with less variability in stimuli.

Furthermore, the production tasks used may have provided less than ideal conditions for any improvements to manifest themselves. More specifically, Hu et al. (2022) found that studies that assessed production with activities involving more spontaneous speech yielded significant effect sizes, while those with activities eliciting more controlled speech did not. Although this is contrary to previous meta-analyses (e.g. Lee et al., 2015), who found larger effects in studies with tasks involving controlled speech, this finding suggests that larger improvements may have been observed if tasks included more spontaneous speech.

Finally, one notable finding concerns CYG learners' use of duration as a cue in L2 vowel perception and production. Even though reliance on durational cues in the perception of the members of a contrastive pair was not examined in this study, these learners did not seem to rely on duration to differentiate contrastive vowels in their productions. This is supported by the fact that duration differences were not consistent across tasks or times, and they were not found in all vowel contrasts. While this is contrary to studies with learners from various L1 backgrounds with fewer monophthongs than the L2 such as Catalan or Spanish (e.g. Aliaga-Garcia & Mora, 2009; Cebrian, 2006, 2007; Kondaurova & Francis, 2010; Mora & Fullana, 2007), Polish (Bogacka, 2004; Rojczyk, 2010) and Chinese Mandarin (Zhi & Li, 2021), similar findings in support of the "full access" hypothesis (Flege & Bohn, 2021) are reported by Lengeris (2009a, 2009b) and Georgiou (2019) for adult SMG and young CYG learners, respectively, who report that learners in their studies did not merely rely on durational cues, but rather used both spectral and durational cues in the perception and categorisation of English vowels, even though neither is used in the L1. Therefore, the results of this study corroborate previous findings for these groups of learners, extending them to production as well, although it remains unclear why SMG and CYG learners behave differently to learners from L1 backgrounds with similar vowel systems.

The following section will focus on a discussion of RQ 4 "How well does any improvement generalise to new speakers and contexts, and to what extent is it retained after a two-month period?". This part of the discussion is divided into two subsections: 5.3.1 focuses on the extent of generalisation of perceptual and production gains after the training, while 5.3.2 discusses the findings in relation to the retention of knowledge as emerged through the delayed post-test (T3).

5.3 Generalisation and Retention of Learning

5.3.1 Generalisation

In perception, generalisation was assessed using three tests, i.e. participants' identification of vowels produced by known speakers in new contexts (/sVt/ and /dVt/), by new speakers in known contexts, and by new speakers in new contexts. In production, generalisation is discussed through participants' intelligibility scores in stimuli produced in new contexts. Due to the limited number of stimuli included in the generalisation tests for each target vowel, the generalisation analyses focused on overall identification and intelligibility scores across

vowels, without analysing individual participant or individual vowel scores. Furthermore, since there was a lack of improvement in the overall production scores of participants across time, no further analyses were conducted on formant values or duration.

5.3.1.1 Perception

In the examination of whether participants were able to generalise improvement to new speakers and/or new contexts, it was established that in all three generalisation tests there were no differences across time (T2 and T3) for either group. Furthermore, while trainees outperformed controls in all tests and both times, this difference was only significant in the test that included new speakers at T3. Importantly, while controls' performance was similar in known and new contexts, trainees were found to perceive vowels significantly better in trained (/bVt/, /gVt/) than in untrained (/sVt/, /dVt/) contexts at both T2 and T3, indicating that the training was effective in improving perception in trained contexts only. However, their T2 and T3 performance was not significantly different than their respective performance in the generalisation tests that involved new speakers.

Overall, the results demonstrated that the performance of learners in either group was not affected by the inclusion of new speakers, but it was when the target vowels were embedded in new contexts. Indeed, a visual examination of the scores shows that both groups performed better in the New Speakers test than their respective performance at T2 and T3 tests, indicating no negative effect of new voices, as opposed to new contexts. It should be noted that the significant differences found in the New Speakers and the New Speakers and Contexts tests should be interpreted with caution, due to the limited number of stimuli included in these tests (36 and 25, respectively). Furthermore, while the two speakers whose voices were included in these tests were chosen randomly, it is possible that some individual characteristics in their pronunciation may have made them more intelligible, which may explain why learners performed better in these tests. A similar observation was made by Lengeris (2009a), where learners performed better in the test with a novel speaker and in an untrained noise condition, leading the researcher to hypothesise that the new speaker may have been more intelligible.

Finally, it should be noted that in an attempt to limit the number of tasks administered to participants as much as possible, the pre-test did not include a perceptual generalisation test, similarly to other studies (e.g. Carlet & Cebrian, 2014). Therefore, based on the data

collected, it is not possible to establish whether participants' performance at T2 and T3 in the generalisation test included any improvement from a pre-test state, as was the case in known contexts and voices, where both groups improved at T2 and T3 compared to T1, but not between T2 and T3.

5.3.1.2 Intelligibility

In production, trainees did not perform differently across the three time points in either task when producing the target vowels in new contexts, similarly to trained contexts. In other words, the lack of an improvement in intelligibility in known contexts was carried over to their productions in new contexts. Furthermore, the intelligibility of trainees' productions was not affected by whether the target vowels were produced in trained or new contexts in either task or time. As opposed to this, controls' productions of the target vowels in new contexts were significantly less intelligible at T3 compared to T1 in the wordlist-reading task. Furthermore, controls performed significantly worse in known than new contexts in the wordlist-reading task at T2. Albeit only significant in this case, learners' productions in untrained contexts were numerically more intelligible than in trained contexts across groups, tasks or times (see **Figure 25** and **Figure 27**). Based on this, it can be hypothesised that the articulatory configurations for the consonants /s/ and /d/ included in the New Contexts test may have facilitated the production of the target vowels by learners, or their perception by NE raters. This is contrary to learners' perceptual patterns, where identification in untrained contexts was more challenging, at least for trainees.

The two groups also performed similarly between them across task and time in new contexts, with the exception of the wordlist-reading task at T3, where trainees outperformed controls. This is similar to their intelligibility scores in known contexts, where trainees outperformed controls in the wordlist-reading task at both T2 and T3. Finally, all participants tended to perform better in the wordlist-reading task compared to the elicitation task in new contexts, with differences reaching significance in trainees for T2 and T3, and in controls for T2, which is a similar pattern to what was found for known contexts as well.

5.3.1.3 General Discussion: Generalisation

Exposure to highly variable input is argued to encourage learners to form more generalised representations of a sound, as it helps them identify relevant phonetic cues and exclude irrelevant, speaker-identity or context-dependent cues. This helps learners develop a more

native-like cue weighting (Aliaga-Garcia & Mora, 2009; Carlet & Cebrian, 2014; Giannakopoulou et al., 2017; Thomson, 2011) and demonstrates that robust learning has taken place (Qian et al., 2018).

In the present study, the lack of a negative effect of new voices on the perceptual performance of participants and the fact that participants in both groups carried over any improvement achieved at T2 and T3 to the generalisation test that included new speakers are taken as preliminary indication that any perceptual improvement achieved by the training can be generalisable to untrained voices. On the other hand, trainees' performance was negatively affected by the inclusion of new contexts at both T2 and T3, suggesting that generalisation of learning to untrained contexts has not taken place, even though measuring learners' T1 performance in new contexts might have revealed some improvement at T2 and T3.

These results are in line with Qian et al. (2018), who found that while their participants improved in their ability to generalise their perceptual discrimination and identification abilities to untrained voices, they failed to do so in new words, suggesting that the training was not successful in facilitating transfer of perceptual gains from trained to untrained words. According to the researchers, longer training sessions and more enriched training stimuli could have been more effective in achieving generalisation to new contexts. Iverson et al. (2005) also report partial generalisation of knowledge to some new contexts but not others in the perceptual identification scores of their participants.

On the other hand, this finding contradicts studies that have observed generalisation of perceptual gains to both new speakers and contexts (e.g. Carlet & Cebrian, 2014; Huensch & Tremblay, 2015; Lively et al., 1993; Wang & Munro, 2004; Zhang et al., 2021); however, it is not unexpected, since previous studies show great variation in terms of the extent of generalisation, despite a general tendency for positive results. For instance, as previously mentioned (5.2.5), variability in stimuli may hinder learning when the target L2 contrasts are highly confusable (e.g. Giannakopoulou et al., 2017; Wade et al., 2007) or when the target group involves perceptually weak or novice learners (e.g. Antoniou & Wong, 2015; Chang & Bowles, 2015; Perrachione et al., 2011; Sadakata & McQueen, 2014), due to the added processing costs required. Although previous studies strongly support that speaker variability in HVPT may promote generalisation, choosing between multiple-talker or single-talker training should take these factors into consideration, especially since some studies have

demonstrated that single-talker training may also be effective in the generalisation of knowledge (Brekelmans et al., 2022; Dong et al., 2019; Perrachione et al., 2011; Wong, 2012, 2014).

As previous studies demonstrate, vowel perception can be strongly affected by consonantal context. For instance, Thomson (2011) found that Mandarin learners of Canadian English were able to improve in the production of vowels in trained contexts (/bV/ and /pV/) and some new phonetic environments (/zV/ and /sV/), but not others (/gV/ and /kV/). The perception of different vowels may also be affected by consonantal context to differing degrees. For example, Bohn and Steinlen (2003) found that the perception of some vowels was more strongly affected by consonantal context, while others were not affected at all. This may also explain the tendency for vowels produced in the untrained contexts to be perceived as more intelligible by NE raters. Learner experience can also determine whether contextual variation affects perception. For instance, Levy and Strange (2008) found that vowels in a bilabial context /bVp/ were more accurately discriminated by American English learners of French than vowels in an alveolar context (/dVt/), but only in inexperienced learners, since more experienced learners did not demonstrate any context effects.

In production, learners' overall intelligibility in new contexts was similar to their performance in known contexts. However, since there was no significant improvement at T2 in the intelligibility scores of participants in known contexts, it cannot be confidently argued that generalisation to new contexts would occur. The findings of this study can merely suggest that context was not a factor that affected trainees' intelligibility in most cases across groups, tasks and times. Therefore, this study can neither support nor contradict previous findings that report a generalisation of knowledge beyond the trained stimuli or voices in the production performance of learners (e.g. Thomson, 2011).

5.3.2 Retention

In this section the findings in relation to the retention of knowledge as measured through the delayed post-test (T3) will be discussed. Participants' performance is discussed based on their perceptual identification scores in trained contexts and voices, and their intelligibility ratings in the production of vowels in trained contexts. Since there was a lack of improvement in the overall production scores of participants across time, no further analyses were conducted on formant values or duration.

5.3.2.1 Perception

Individual participant performance at T3 ranged from poor to good (26.4% to 68.9%), similarly to T2. All participants who improved at T2 maintained or increased their performance at T3 as well (CYG03, CYG04, CYG08, CYG12, CYG15 and CYG17). However, one trainee (CYG01) and one control (CYG06) showed improvement at T3 only. Some of these learners demonstrated increasing improvement across the three time points (CYG01, CYG03, CYG04 and CYG06), while others performed similarly between T2 and T3, but still better than at T1 (CYG08, CYG12, CYG15 and CYG17). Six of these participants belonged to the trainee group (CYG01, CYG03, CYG04, CYG08, CYG12, CYG17), while two were controls (CYG06 and CYG15). However, since different patterns were observed in different participants, with some of them showing decreased performance at T2 or T3 compared to T1, these improvements cannot be attributed to familiarity effects. Group differences were observed only at T3, where trainees significantly outperformed controls. Furthermore, while both groups improved from T1 to T2 and from T1 to T3, no improvement was observed between T2 to T3 in either group. This shows that any improvement from T1 to T2 was retained, but in both groups. Overall perception remained moderate (56.3% for trainees, 41.2% for controls).

Finally, in individual vowels, the improvement observed in trainees' performance in KIT, FLEECE, BATH, TRAP and NORTH at T2 was maintained at T3 as well, while NURSE and GOOSE were also improved in perception between T1 and T3, but not between T1 and T2. The improvement observed in KIT in the control group at T2 was also maintained at T3, but no other vowels improved significantly. Differences in the identification of target vowels between the two groups at T3 were significant in FLEECE, NURSE, BATH, NORTH and GOOSE. Overall, perception ranged from poor (16.7% for NURSE in controls) to very good (82% for GOOSE in trainees) depending on the vowel. Again, these improvements cannot be attributed to familiarity effects, since different patterns were observed in different vowels, with some of them having lower identification scores at T2 or T3 compared to T1, although insignificant.

5.3.2.2 Intelligibility

In individual learners, significant differences in intelligibility scores at T3 compared to T1 were observed only in three trainees (CYG03 in the elicitation task, and CYG12 and CYG16 in the wordlist-reading task). The former two participants had lower intelligibility scores at T3

compared to T1, even though no differences were observed between T1 and T2. The latter had a significantly higher intelligibility at T2, which was maintained at T3 as well.

The comparison between the two groups of CYG learners at T3 showed that while they performed similarly in the elicitation task, trainees performed better than controls in the wordlist-reading task. The overall performance of participants in both tasks showed that while the two groups performed similarly at T1, trainees received higher intelligibility scores than controls at T3, similarly to T2. It was further established that these differences were found in the wordlist-reading task (similarly to T2), since the two groups performed similarly in the elicitation task across time. However, no effect of time was found in the performance of trainees in either task, and therefore, it can be concluded that the training was not adequate to cause any changes in the overall performance of this group across the three time points.

In individual vowels, only two vowels improved from T1 to T2 for trainees: BATH in the elicitation task and FLEECE in the wordlist-reading task. While this improvement was not retained for BATH, as indicated by the similar values observed between T1 and T3, the improvement was retained for FLEECE, as suggested by the higher intelligibility score at T3 compared to T1. A different pattern was observed for trainees' productions of FLEECE and LOT in the elicitation task. More specifically, these vowels became significantly less intelligible at T2 compared to T1, but their intelligibility increased again at T3, where no significant differences were found compared to T1 scores. As noted in 5.2.3.2, this inconsistency across different vowels suggests that learners were possibly overwhelmed and confused upon realising the existence of different L2 vowels, which possibly led to unsuccessful modifications of their productions in an attempt to differentiate them.

5.3.2.3 General Discussion: Retention

As mentioned in 2.8.3, assessing retention enables the examination of whether learners have learnt to turn the muscular control required for the production of L2 segments into a habit. According to Koutsoudas and Koutsoudas (1962), learning to perceive and produce sounds with conscious effort without acquiring the new habit may not lead to automatic production, since in this case learners are likely to continue to use familiar habits in the production of L2 sounds. Therefore, it is important to examine whether learners' performance is turned into a habit, which continues to be used long after the training.

Overall, all participants who improved in perception at T2 maintained or increased their performance at T3 as well, while one trainee and one control improved at T3 only. Furthermore, both groups improved their perceptual performance at T2 and maintained this improvement at T3 as well. The improved perception of some vowels at T2 was also maintained or increased in the same vowels at T3. To these, NURSE and GOOSE were added for trainees, who perceived these vowels better at T3 compared to T1, even though no significant improvements were observed between T1 and T2. These results demonstrate that the improvements observed in perceptual performance were retained two months after the training.

In production, only one trainee performed significantly better in intelligibility at T2, and his increased performance was maintained at T3. Trainees outperformed controls in intelligibility at T2, which was maintained at T3, but further analyses showed that this was only in the wordlist-reading task. However, it should be noted that trainees did not improve significantly as an effect of time. Based on these results, retention could not be established with confidence in the production performance of participants in this study. While the fact that one trainee produced more intelligible vowels and was able to retain this performance in the delayed post-test is encouraging, there was no significant improvement in either group's performance across time, which made it impossible to examine retention.

Overall, these results demonstrate that despite the fact that perceptual or production improvements were very few in the study, and limited to certain participants, tasks or vowels, these improvements were to a large extend retained two months after the training, with the exception of the improvement of BATH in trainees' productions in the elicitation task. This retention of learning is in agreement with findings of previous research regarding segmental perception and production (e.g. Carlet, 2019; Cebrian et al., 2019; Rato & Rauber, 2015; Thomson, 2012; Wang & Munro, 2004), where learners were found to retain improvements for periods ranging from one to four months later.

Importantly, the results of this study indicate that additional input received from other sources after the post-test may have promoted perceptual learning further through more conscious (or unconscious) effort paid by the participants, especially trainees, in the perception of the target vowels. This may be supported by three observations: firstly, the improved perceptual performance of two learners, which reached significance only at T3;

secondly, the fact that trainees' performance was significantly better than controls' performance only at T3; and thirdly, the observation that two additional vowels were perceived better by trainees at T3.

Another important result concerns some participants' poorer performance at T2 compared to T1 or T3 (CYG03, CYG05, CYG12 and CYG15). It is hypothesised that this pattern may be due to confusion arising after awareness of the existence of contrasts in L2 vowels, in combination with a lack of adequate input that could help learners identify and produce them using the relevant acoustic cues. This may be explained by Flege and Bohn's (2021) observation that the time course of L2 category formation is not fully understood yet, and it is possible for learners to experiment with their productions before reaching more accurate configurations at a later stage when additional input is received.

In terms of the perception-production link, it was observed that perceptual and production results were not aligned in any comparison. In terms of group performance, while both groups improved in perception from T1 to T2 and T3, trainees did not improve in intelligibility across time in either task, while controls performed worse at T2 compared to T1 or T3 in the elicitation task. It should also be noted that the perceptual improvements observed by both groups were not reflected in their productions in the elicitation task, despite the fact that this task entails an element of perception (Nagle & Baese-Berk, 2022). More specifically, the two groups had similar intelligibility scores in this task, despite the fact that trainees demonstrated a better perceptual performance than controls at T3. Furthermore, while both groups improved at T3 compared to their T1 perceptual scores, their intelligibility scores remained similar between the two times in both tasks.

In individual participant performance, only one participant (CYG16) had improved intelligibility scores at T2 which were maintained at T3 as well, but he showed no improvement in perception at any time. At the same time, other participants who showed perceptual improvement at T2 and/or T3 (CYG01, CYG03, CYG04, CYG06, CYG08, CYG12, CYG15 and CYG17) did not demonstrate any improvement in intelligibility. Finally, while the perception of several individual vowels was improved for trainees between T1 and T3 (KIT, FLEECE, BATH, TRAP, NORTH, NURSE and GOOSE), as was the perception of KIT for controls across time, none of the vowels were significantly improved in intelligibility between T1 and

T3. The only improvement observed in intelligibility at T2 that was retained at T3 as well was in FLEECE produced by trainees in the wordlist-reading task.

However, these results should be interpreted with caution for two main reasons. Firstly, in perception, one participant in the control group improved and retained this improvement at T3 (CYG15), while another demonstrated improvement at T3 only (CYG06), suggesting that other factors may have also played a role in participants' performance. Secondly, perceptual improvement was observed in both groups at T2, and this was retained by both groups at T3, as indicated by a similar performance between the immediate and delayed post-test, both of which were significantly improved compared to the pre-test. This finding contradicts Carlet (2019), who reported similar results but only for the experimental group; the researcher argued that the fact that the control group performed similarly across time was an indication that robust learning had taken place for the experimental group. In the present study, the fact that both groups demonstrated the same pattern weakens the argument that robust learning has taken place for the trained group.

Having examined CYG learners' perceptual and production patterns across time and in known and new contexts, the final part of the discussion turns to an examination of the effects of individual differences, in order to address RQ 5 "Do individual differences in motivation, input and language use patterns affect CYG learners' perception and production of L2 English vowels?".

5.4 Individual Differences in Perception and Production

The fact that trainees performed better than controls in most tests, despite the lack of a notable effect of the training, suggests that other factors may have played a role in their performance. Many researchers (e.g. Cebrian et al., 2019; Melnik-Leroy et al., 2022; Nagle & Baese-Berk, 2022; Piske et al., 2001; Rato & Carlet, 2020, among others) suggest examining the performance of individual learners to better understand how L2 segments are perceived and produced. For instance, it is reasonable to expect that learners who are more motivated to improve their pronunciation or who use the L2 more frequently will be more likely to develop both modalities further than learners who lack motivation or use the L2 less often (Gilakjani, 2012; Nagle & Baese-Berk, 2022). This section explores the effects of motivation, input and language use patterns on individual learner performance using both the quantitative and qualitative data collected.

5.4.1 Motivation

Quantitative analyses showed that motivation had a significant positive correlation with participants' intelligibility scores in both tasks at T1, but not T2 or T3. Motivation was also positively correlated with input received during their university years and total input, as calculated in 4.5.2. However, no correlation was found between participants' motivation scores and their perceptual performance at any time. Qualitative analyses further supported this finding. More specifically, the two participants who demonstrated a high production performance at T1 (CYG03 and CYG08) strongly expressed their motivation to try and improve their pronunciation and stated that they use English more, both with friends and relatives, taking every opportunity they have to use English and practice their pronunciation skills. On the other hand, CYG01 expressed a lack of motivation to improve his pronunciation skills, and apart from his lectures, he reported little interaction in English with family and friends.

These results are in line with several previous studies that observed a positive correlation between motivation and pronunciation performance (e.g. Elliott, 1995; Moyer, 2007; Purcell & Suter, 1980; Suter, 1976; Yousofi & Naderifarjad, 2015). Participants' responses also support Gilakjani's (2012) assumption that L2 learners who are concerned about their pronunciation are more likely to perform better, at least partly because they are more likely to seek out opportunities to use and improve their skills in the L2, as was reported by participants CYG03 and CYG08. On the other hand, some learners may not have the motivation to improve their performance, even if aware that their speech is foreign-accented, as long as their ability to communicate effectively is not compromised (Moyer, 2007), as was the case for CYG01.

Furthermore, all three interviewees in the present research expressed their belief that a good pronunciation can have a positive effect in their professional life, but felt that their pronunciation level is adequate to meet employers or lecturers' expectations. However, interviewees had differences in terms of their personal and professional goals, which were reflected in their production performance. For instance, CYG03, who had the highest intelligibility scores at T1, expressed the most positive attitudes towards the L2, personal motivation for self-improvement, as well as the intention to study abroad, demonstrating both instrumental (i.e. referring to linguistic achievement) and integrative (i.e. referring to a desire to learn about the culture of the target language) motivation (Gardner & Lambert, 1972). On

the other end, CYG01, who had one of the lowest intelligibility scores at T1, expressed no interest in improving his pronunciation further than necessary for communication, and no immediate goals that would require a high level of pronunciation skills. This relationship is in agreement with previous literature. For example, Moyer (2007) found that a desire to improve accent, a positive orientation towards the L2 and the L2 culture, comfort with assimilation and having the intention to reside long-term in an L2-speaking country can be important factors in improving pronunciation, while Gilakjani (2012) noted that personal and professional motivation to learn English can have a positive effect on the desire of learners to reach native-like pronunciation.

One additional observation related to the fear of stereotyping or anticipated stigmatisation in a professional context as a result of foreign-accented speech, which was expressed by participant CYG08. Despite her concern, however, the participant also stated her belief that achieving communicative competence in the L2 is adequate to overcome negative evaluations. This statement is supported by previous studies as well. More specifically, although it has been observed in previous studies (e.g. Baquiran & Nicolaidis, 2020; Buckingham, 2014; Hendriks et al., 2016, 2018, 2021) that foreign-accented speakers may face bias as an effect of their accent in terms of perceived competence, comprehensibility and likeability, among other characteristics, some studies showed that non-native speakers with a slight accent are viewed as similar to native speakers (e.g. Carlson & McHenry, 2006; Dragojevic et al., 2017; Hendriks et al., 2021; Said, 2006), which demonstrates that strength of foreign accent plays a role in how non-native speakers are evaluated, with stronger accents eliciting more negative evaluations.

On the other hand, the results of the present study contradict some previous findings, such as those of Thompson (1991), who found no significant effect of motivation on foreign accent. As noted in 2.6.2, the contradictory findings reported in such studies are likely due to the fact that motivation is difficult to measure precisely and consistently across studies, as well as the fact that studies assessing the effect of motivation rely on participants' self-ratings, which makes it difficult to know the accuracy of their responses (Piske et al., 2001). Although an attempt was made in the present study to mitigate this limitation by the use of both quantitative and qualitative data, the analyses can merely confirm participants' consistency in their responses but not their accuracy.

Furthermore, differences among learners may arise from a combination of other factors coexisting or confounded with motivation as well, such as the quality and quantity of L2 input, age of first exposure and L1-L2 use patterns, which also contributes to the inconsistencies observed in previous studies (Piske et al., 2001). For instance, as shown by both the quantitative and qualitative data collected in the current study, participants' motivation was associated with frequency of interactions with native and non-native speakers of English, input received, as well as willingness to practice L2 skills, which makes it difficult to disentangle the effects of motivation from those of other factors.

In summary, the results of this study add to the previous literature that has found at least a small effect of motivation, indicating that this factor should not be neglected in future research and should be taken into consideration in L2 pronunciation instruction. The fact that none of the participants reached native-like competence, despite the high degree of motivation of at least some participants, cannot be used as an argument against the significance of the role of motivation. A similar finding is reported by Moyer (1999), who observed that despite a strong correlation between professional motivation and foreign accent, none of the participants reached pronunciation ratings similar to those of native speakers. This is also in line with Wells' (2005) argument that motivation alone may not be enough to lead to native-like proficiency. Importantly though, the present research does not argue that motivation will lead to native-likeness, but rather that it can contribute to an improved performance for some highly-motivated learners compared to their less motivated peers.

5.4.2 Input and L1-L2 Use Patterns

Quantitative analyses showed no significant correlations between participants' intelligibility scores and input or L1-L2 use patterns. A significant positive correlation was only observed between participants' perceptual performance at T1 and total input received, i.e. the combined input received during their school and university years. This correlation was not found in perceptual performance at T2 or T3. The lack of a significant effect of input in most cases in the present study is in line with some previous studies (e.g. Cebrian, 2006; Flege & Wayland, 2019), and seems to support the argument that input only has a limited role to play in pronunciation, as suggested by some researchers (e.g. DeKeyser & Larson-Hall, 2005; DeKeyser, 2000).

One possible explanation offered by Flege and Wayland (2019) is in support of Lenneberg's (1967) initial formulation of the Critical Period Hypothesis, which suggests that learners cannot automatically use input by mere exposure to it and that the success of some late learners may be due to conscious effort. According to this hypothesis, successful late learners may reach native-like production only because of a special aptitude for L2 learning or of strong motivation that led them to work hard to learn the L2 pronunciation. A second explanation offered by Flege and Wayland (2019) relates to the continued influence of the L1, in combination with an inadequate amount of L2 input received. While the results of the present research seem to be in line with the first explanation, Flege and Wayland (2019) warn that choosing between the two explanations necessitates the identification of accurate methods to measure quantity and quality of input. While the present research attempted to use qualitative interviews to gain more insights into the quality and quantity of input participants received, it still relied on the self-estimates of participants on overall L2 use, including recalling past habits, which remains an inadequate approach to reliably measure amount of input as well as how much of it is foreign-accented (Flege & Wayland, 2019; Flege, 2008).

Despite this limitation, the qualitative analyses included in this study provided some further insights into the complex nature of the role of input. More specifically, the three participants reported a similar classroom experience during their school and university years, although they differed in their use of English outside the classroom. The linguistic practices of the three participants can be divided into two categories, forming an interesting pattern: interviewees CYG03 and CYG08 reported a higher degree of interaction in English with native speakers of different varieties and non-native speakers, which includes both speaking and listening to the L2, as well as conscious efforts to practice their pronunciation, and achieved two of the highest intelligibility scores at T1, outperforming CYG01; on the other hand, CYG01 reported extensive movie-watching, which only involves listening to the L2, and outperformed both CYG03 and CYG08 in perceptual performance, achieving one of the highest, albeit moderate, perceptual scores at T1. It is therefore hypothesised that the passive nature of movie-watching may have helped this participant in perceiving the target segments, but not in producing them, since no conscious effort was made to improve production, as reported in his interview. Although this analysis was based on very few learners and cannot be used as

robust evidence of the effect of input, it demonstrates a pattern that merits further attention in future research.

Importantly, as Flege (2008) points out, both quantity and quality of input should be taken into consideration when examining the effect of this factor, since L2 input is different and more variable compared to L1 input. CYG learners' performance in both perception and production was likely affected by the quality of the input received, which did not necessarily reflect the target variety of the present study. For instance, the relatives reported by CYG03 and CYG08 are bilinguals in CYG and English, meaning that the input received was likely different compared to the monolingual SSBE speakers included in this study. Furthermore, as per their reports, these relatives come from various regional backgrounds, meaning that the input received varied for each learner. At the same time, CYG01 reported watching movies that included a variety of accents. Furthermore, all participants reported extensive exposure to non-native speaker input, including their CYG teachers and classmates at school and university.

As noted by Flege and Wayland (2019), exposure to different input distributions is likely to lead to individual differences in the phonetic categories developed for each learner, which depend on the acoustic properties present in the input encountered. Furthermore, Flege and Liu (2001) argued that L2 speakers may learn to accurately perceive and produce L2 sounds, but only if they receive adequate and high-quality native-like input. Therefore, the different types of input each participant received may explain individual differences in their performance, as well as the fact that their overall perceptual and production performance did not reach native-like values.

The effects of the input received from GA should also be noted, since it is possible that CYG learners were exposed to this variety more than SSBE, similarly to the Serbian students in Čubrović and Bjelaković (2020), due to the influence of movies, music and the media. Importantly, Čubrović and Bjelaković (2020) found that Serbian students used a mixture of features from both varieties, with only a few learners showing consistency in their use of specific pronunciation features. The researchers explained this as a result of the exposure to GA in combination with instruction in SSBE, which had possibly led to the mixing of features of the two varieties in the speech of L2 learners. Although the effect of GA input was not examined in the present research, it is likely that CYG learners follow the same pattern in their

perception and production of L2 English vowels. This is also in line with Zhi and Li (2021), who noted that L2 speakers acquiring English in non-English-speaking countries are exposed to different English accents by their teachers, textbook recordings, the media and other sources, and are therefore likely to develop mixed accents in their L2 speech.

Finally, according to Nagle and Baese-Berk (2022), individual differences in language use patterns may also influence the outcomes of pronunciation training. In the current study, interviewees' intelligibility scores remained mostly unaffected by the training, except for CYG03 who had a lower intelligibility score in the elicitation task at T3 compared to T1. On the other hand, the perceptual scores of all three interviewees improved after the training, although for CYG01, the improvement reached significance only at T3. Furthermore, while CYG01 outperformed CYG03 and CYG08 in perception at T1, CYG03 and CYG08 achieved similar or higher perceptual performance than CYG01 at T2 and T3. This supports Nagle and Baese-Berk's (2022) argument that learners who use the L2 more frequently and extensively (such as CYG03 and CYG08) are more likely to develop both modalities and the link between them, compared to learners who use the L2 less often.

In summary, although quantitative analyses demonstrated a weak effect of input and language use patterns on the perception and production of L2 segments, qualitative analyses showed that this factor is far more complex than what can be captured by a questionnaire. As the results of this study show, and as noted by Tyler (2019), L2 learners in an EFL environment vary in many respects, including the amount of previous experience they have with the L2 and their exposure to different regional and foreign accents, which is a factor that should be taken into consideration when developing L2 teaching curricula.

The discussion of the effects of individual differences in the perception and production of L2 segments concludes this chapter. Having discussed the findings of the present study for each RQ and in relation to previous studies, the final chapter of this thesis presents some concluding remarks. The first part of the chapter states some limitations and directions for future research; finally, some summarising and concluding remarks of the study are provided, which also highlight its contribution to the field.

CHAPTER 6: CONCLUDING REMARKS

6.1 Limitations and Further Research

The present study suffers from various limitations that should be pointed out and could provide directions for future research. Firstly, in choosing the stimuli for the tasks of the study, priority was given to ensuring that the target vowels appeared in matching environments to the extent possible over considering word-frequency or word-familiarity effects. This was considered as the best option for two main reasons: firstly, potential word-frequency effects were minimised by the inclusion of both real and non-words (Huensch & Tremblay, 2015); secondly, some researchers (e.g. Carlet & Cebrian, 2014; Thomson & Derwing, 2016) have previously noted that while lexical frequency or familiarity may have an effect, it was not found to significantly impact identification or intelligibility scores in their studies. However, since the present research did not examine this, future studies could incorporate an examination of potential word-frequency and word-familiarity effects to confirm these assumptions or reveal further insights.

In addition, the inclusion of both real and non-words to ensure matching environments for the target vowels as well as forcing learners to attend to relevant phonetic details rather than meaning also entails limitations. Previous studies report contradictory results in terms of the perception of L2 categories in real or non-words. For instance, Thomson and Derwing (2016) found that phonetic training with predominantly non-words led to larger improvements than using predominantly real words, while Cebrian et al. (2019) and Rato and Carlet (2020) reported that perceptual performance was better in real words than in non-words, and that the perception of different target vowels was affected differently depending on stimulus type. Even though a visual inspection of **Table 4** shows no consistent pattern in the identification of target vowels in real vs. non-words, this study did not examine learners' performance in the two types of stimuli in detail, firstly because it was not part of the main aims of this research, but mainly due to the small sample size that would emerge by analysing performance in the two types of stimuli separately, which would not yield any robust conclusions. Further research should take this factor into consideration so that more conclusive results can be reached in order to establish best practices in training designs.

Another limitation concerns the use of a /gVd/ or /gVld/ rather than a /gVt/ context for both vowels in the FOOT-GOOSE contrast, which was preferred in spite of the priority given to

maintaining matching contexts mainly to ensure that the target vowels were as orthographically unambiguous as possible to both NE speakers and CYG learners, so that the vowels were produced as intended. This may have affected the perceptual identification of the target vowels in this contrast. For instance, the high perceptual identification scores of the two vowels by CYG learners could arguably be explained by the presence of the post-vocalic /l/ in the stimulus word for GOOSE, i.e. “gould”, and its effect on the quality of the preceding vowel, which may have facilitated the identification of the target vowel in this pair. Indeed, **Table 4** shows that whereas FOOT was not very well identified in the /bVt/ context, the two contrastive vowels were identified at a much higher percentage in the /gVt/ context. However, it should be noted that a similar pattern is observed in the TRAP-STRUT contrast as well, where the three stimuli “bat”, “gat” and “but” were identified at a similar rate, while “gut” was much more challenging for learners despite the fact that the same environment and all real words were used. This suggests that the effect of this limitation may have been smaller than anticipated, although future studies should consider alternative ways to ensure that orthographic representations match the intended phonetic realisations without altering phonetic context.

The limited number of stimuli for each target vowel in the generalisation tests should also be mentioned. Although this choice was made to assess generalisability of learning to real words only, this smaller sample size prevented the analysis of participants’ performance in each target vowel separately and led to a lower statistical power of the analyses conducted. In addition, learners’ better performance in the New Speakers test may have been due to the smaller sample size; an alternative possibility is that some individual characteristics in the pronunciation of the chosen speakers may have made them more intelligible. Therefore, further research should be conducted to confirm the findings of the study or provide additional insights.

Furthermore, although the HVPT paradigm used was carefully considered based on the efficacy of practices as observed in previous studies, and on practicality, so as it would not be discouraging to participants or interested parties, it may also entail some limitations, especially since there is great variation in the paradigms used in previous research and a lack of agreement as to what is considered as best practice for optimal learning (Thomson, 2018). This makes it difficult to choose a specific approach and to make comparisons with other

studies or identify which specific aspect or combination of aspects was effective in promoting gains. While it was not the aim of this study to assess or establish a set of best practices for HVPT, the lack of a clear effect of the training on participants' perception and production performance is likely due to methodological choices. For instance, the length of the training may have been inadequate in relation to the number of target segments and the degree of variability in training stimuli may have been counterproductive. Therefore, the effects of the protocol followed can add to the existing literature and guide future research, which can compare different methodological approaches to reach conclusions as to the most effective techniques to be incorporated into HVPT to promote maximal gain, but also to ensure that the findings of different studies are comparable.

Incorporating explicit production practice during perceptual training has been suggested in previous studies (e.g. Thomson & Derwing, 2016) and was also considered, but it was avoided for several reasons. Firstly, it would be impractical, as it would require the completion of additional sessions by the participants, which would likely be discouraging; other tasks, such as the generalisation and retention tests were deemed more useful to include at the expense of a production component based on the aims of the study. Secondly, one of the aims of this research was to examine whether training in one modality could be beneficial for both, so as to lead to recommendations for L2 practice that would assist rather than burden language teachers. Last but not least, previous research has yielded contradictory findings as to whether incorporating production is beneficial or detrimental to learning. For instance, while Wong (2013) found that high-variability perception training when combined with production training was more effective than training on each modality alone, Nagle and Baese-Berk (2022) note some studies (i.e. Baese-Berk & Samuel, 2016; Baese-Berk, 2019) which found that producing sounds during perceptual training can have disruptive effects on perceptual learning.

In addition, this study included the examination of factors that are difficult to measure reliably, i.e. motivation, input and L1-L2 use patterns. Although attempts were made to mitigate this limitation through the completion of a detailed questionnaire as suggested by Derwing and Munro (2015) and the use of qualitative interviews as suggested by Moyer (2008), the examination of these factors still relied on the self-estimates of participants and in some cases, on their recollection of past habits, an approach that does not provide an

objective measure of the variables in question. Future studies should consider incorporating more reliable and objective measures of these variables; one such method for measuring quality and quantity of input is suggested by Flege and Wayland (2019) and involves asking participants to respond to notifications asking them to record stimuli as well as information about their interlocutors at specific moments over a period of time.

Another limitation concerns the examination of the link between perception and production. Nagle and Baese-Berk (2022) suggest that research on the link between the two modalities should focus more on how it develops and changes over time. Since perceptual learning is developmental in nature, it should be examined longitudinally and not through a single test at a single point in time (Nagle & Baese-Berk, 2022). The current research examined participants' perceptual and production performance at specific time points within a period of only two months, and as a result, it cannot provide further information about this developmental nature of the link between them or any changes across time. Therefore, as the researchers suggest, future studies could incorporate a variety of perception tasks at various time points in order to further our understanding of L2 perceptual learning.

Another consideration for future studies could be to compare L2 learners not only to monolingual native speakers as in this study, but also to advanced L2 users that are highly intelligible and comprehensible (Nagle & Baese-Berk, 2022). This is because monolingual norms are different, and given the fact that the aim of such research is to achieve intelligible and not necessarily native-like speech, comparisons with monolinguals may not reveal the complex ways in which L2 learners perceive and produce L2 contrasts (Nagle & Baese-Berk, 2022). At the same time, research in the field could include raters from other L1 backgrounds in addition to NE raters, in order to investigate their perceptions as well.

Importantly, as stressed by other researchers (e.g. Baese-Berk et al., 2020; Hu et al., 2022), the role of the listener should not be neglected. Although the vast majority of studies has placed the communicative burden upon the non-native speaker, successful communication relies on the native listener as well, who may bring into the conversation his or her own social and linguistic experiences, familiarity with particular accents (or lack thereof), biases, expectations and cognitive abilities, which play an important role in their understanding and perception of non-native accented speech (Baese-Berk et al., 2020; Dragojevic & Goatley-Soan, 2020; Hu et al., 2022). Despite the importance of understanding

the acoustic properties of non-native speech and how it differs from native speech, both interlocutors are mutually responsible for reaching mutual understanding (Baese-Berk et al., 2020). Therefore, understanding the factors that may impact native listeners' perception of non-native speech is equally important, and future research should also place more focus on the role of the listener as well.

Finally, one limitation that may have affected the results of the study arose from the restrictions imposed during the COVID-19 pandemic, which affected various aspects of this research. For example, participant recruitment was particularly challenging during that period, mainly due to movement restrictions and lack of access and contact with potential participants. Even though some previous HVPT studies (e.g. Kondaurova & Francis, 2010; Lively et al., 1993; Logan et al., 1991) have also reported a limited number of participants per group, perhaps due to their time-consuming nature, the small sample size in the present research was not ideal. In addition, the initial aim to administer all tests and training sessions in face-to-face meetings was not possible, and due to the need to limit interactions as much as possible, only the production tasks in the study were completed in face-to-face sessions. As a result, it was not possible to closely monitor participants' completion of the perceptual tasks and training, despite attempts to mitigate this with the use of online screen-sharing and remote desktop software.

6.2 Summary and Conclusions

The aim of this study was to provide an in-depth examination of the acquisition of L2 English vowels by native speakers of CYG and the factors that may affect their perception and production. It has firstly analysed CYG learners' perception and production of L2 segments before any intervention and compared them to the production patterns of NE speakers evaluating the effects of the L1 as well. The study then assessed the effects of HVPT to examine whether learners were able to shift their patterns towards more native-like norms and evaluated whether any improvements were generalisable to new speakers and contexts or retained two months later. Finally, the effects of motivation, input and L1-L2 use patterns were examined through both quantitative and qualitative data.

Research on the phonetics and phonology of CYG is limited and rarely focuses on vowels, while previous studies examining L2 English vowel acquisition by SMG or CYG learners have focused on the perception modality, paying little attention to production patterns (e.g.

Georgiou, 2019; Lengeris & Hazan, 2007; Lengeris, 2009a, 2009b). Although the vowels of the dialect have been acoustically analysed and compared to SMG vowels (Themistocleous & Logotheti, 2016; Themistocleous, 2017a, 2017b), to my knowledge, no previous study has compared CYG to English vowels to determine differences in their acoustic space. Therefore, this study provides insights into the realisation of vowels in one of the most widely spoken varieties of Greek, which remains largely unstudied, and provides a clearer picture of the production as well as the perception patterns of adult CYG learners in L2 English.

Examining CYG as a separate variety from SMG is important, since substantial variation can be found even in simple five-vowel systems as an effect of sociolinguistic factors and linguistic context (Arvaniti, 2010). Apart from the subtle differences between the vowel inventories of SMG and CYG (2.3.1), the political and historical background of Cyprus as outlined in 2.1, as well as the tendency of Greek-Cypriots to study in English-speaking countries may affect these learners differently, possibly providing additional motivation or affecting their language use patterns. For example, as opposed to SMG learners, and as revealed through the qualitative analysis of this study, CYG learners are more likely to interact in English in other contexts as well, such as with relatives that reside permanently in English-speaking communities or with immigrants living and working in Cyprus, which are very common in Cyprus.

The results of this study clearly showed the influence of the L1 on both the perception and production of L2 segments, supporting the assumptions of current models of speech perception and production such as the PAM-L2 and the SLM-r. CYG learners faced challenges in perceiving the members of an L2 contrast and used their L1 articulatory routines to produce the L2 vowels, irrespective of the acoustic distance between them. This is an important finding, since identifying the difficulties that learners might encounter and how their L1 could influence L2 speech production and perception can guide EFL teachers and can help in the development of EFL curricula.

Furthermore, while previous studies incorporating HVPT with learners from various L1 backgrounds have demonstrated that it is possible for adult L2 learners to improve their perception and in some cases production of L2 sounds given sufficient training, only a handful of HVPT studies included native speakers of SMG (e.g. Giannakopoulou et al., 2017; Lengeris & Hazan, 2010; Lengeris, 2008, 2009a, 2018), and to my knowledge, only Georgiou (2021) has

investigated CYG learners of English. The study conducted by Georgiou (2021) provided some informative preliminary results for this new group of learners; the present research offers further insights with the use of more extensive training and examines the effects of the training in the long-term retention of knowledge as well as investigating production performance in more detail through acoustic analysis in addition to native-speaker judgments.

The examination of the effects of HVPT in the present study showed that the intervention was not sufficient to lead to significant improvements in trainees' overall perceptual or production performance, although some improvements were observed in either perceiving or producing some target vowels. This also restricted the assessment of generalisation and retention of learning, since it was only possible to examine them in those areas where improvement had occurred. Furthermore, the results of this study showed no direct link between perception and production before or after the training, supporting the hypothesis that the two modalities develop independently from one another.

Learners' performance both before and after the training was hardly surprising given the limited and often foreign-accented input they had received and the lack of naturalistic exposure throughout L2 learning. Despite these discouraging results, however, the present study can increase awareness of this technique in Cyprus so as to encourage the development of improved versions of it that can be used by EFL teachers and interested individuals in the L2 classroom or as a tool for self-study, something that has not been achieved so far, despite the modern-day technological advancements. The results of this study may encourage universities in Cyprus or elsewhere to place more focus on HVPT in their course material, enabling non-native students that are likely to become English language instructors to improve their teaching practices, as well as their own pronunciation skills, thereby bridging the gap between research and practice. In addition to teachers, course developers and material writers can also take advantage of these findings to update and improve language pedagogy as well.

Several researchers (e.g. Gilakjani, 2012; Tsang, 2022; Tyler, 2019) have pointed out the need for teachers in the EFL classroom to increase learners' awareness of the importance of pronunciation, provide them with adequate opportunities for practice as well as exposure to input that preserves phonological contrasts between L2 phonemes and guide them in developing intelligible speech within and beyond the classroom. Although accent-free speech

should not necessarily be the ultimate goal, L2 pronunciation instruction should focus on teaching L2 segmentals in a way as to enable learners to better recognise and produce them to achieve their communication goals (Thomson & Derwing, 2016). However, teachers often lack the tools, experience or knowledge to properly teach pronunciation. Therefore, when tested and found to be effective, tools such as the HVPT can provide the required exposure and native-speaker input.

Furthermore, this study has integrated both linguistic factors and individual differences in the examination of CYG learners' perception and production of English vowels. In addition to the influence of the L1 on L2 speech, the individual variability observed in previous studies suggests that L2 production depends on more than just L1-L2 category mappings. The examination of factors affecting degree of foreign accent in this study provides an attempt to account for individual differences in the perceptual and production performance of CYG learners and their interaction with HVPT. The importance of investigating and understanding individual learner differences and their role on the impact of training is also stressed by Barriuso and Hayes-Harb (2018), as HVPT materials and procedures may need to be tailored to individual needs and abilities in order to increase the effectiveness of HVPT, rather than using the same approach for all learners. Nagle and Baese-Berk (2022) also noted the need for research examining individual differences and their effect on each modality.

This study offers insights into the importance of motivation, input and language use patterns, which have not been previously examined in the context of Greece or Cyprus. The results indicated that individual differences among the learners may have indeed affected their performance. Both quantitative and qualitative data demonstrated that motivation can play a significant role in learners' pronunciation performance and that input can positively influence perceptual performance, which means that both factors should be taken into consideration in L2 pronunciation instruction. Assessing the significance of each factor investigated in the present study can guide future research in the development of appropriate strategies and classroom practices related to each factor in order to improve pronunciation instruction, especially for adult EFL learners. Furthermore, the results of this study also highlight the fact that individual differences among L2 learners in an EFL classroom make predictions about L2 segmental acquisition very difficult and therefore they should be taken into consideration when developing EFL curricula (Tyler, 2019).

The overall findings of the present study are particularly important given that little attention is paid to pronunciation in the EFL classroom in Cyprus, despite the negative implications of pronunciation errors for L2 learners, such as negative evaluation, discrimination and speaking anxiety (2.7.2). This is an issue that needs to be addressed, since the ultimate goal of teaching English should be to enable learners to participate in any communication where English is used, and intelligible pronunciation is considered to be essential for communicating effectively (Kyprianou, 2007; Wells, 2005). Finally, it should be noted that although the results of this study could be used at schools with younger learners if future research ascertains that the findings can be applied to children, it is primarily focused on adult learners, i.e. those learners that may wish to reduce their foreign accent in English, perhaps for professional reasons. Using HVPT outside the classroom requires motivation, which may be the case for adults who have specific personal or professional reasons for wanting to improve their intelligibility and reduce their degree of foreign accent. This is possible, since CAPT applications offer an easily accessible and self-paced instruction, that can be used individually or to complement classroom pronunciation teaching.

LIST OF REFERENCES

Akomolafe, S. (2013). The invisible minority: Revisiting the debate on foreign-accented speakers and upward mobility in the workplace. *Journal of Cultural Diversity, 20*(1), 7-14.

Aliaga-Garcia, C. (2017). *The effect of auditory and articulatory phonetic training on the perception and production of L2 vowels by Catalan-Spanish learners of English* (Doctoral dissertation). Universitat de Barcelona. Retrieved from https://diposit.ub.edu/dspace/bitstream/2445/121323/1/CAG_PhD_THESIS.pdf

Aliaga-Garcia, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In M. A. Watkins, A. S. Rauber & B. O. Baptista (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (pp. 2-31). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Alispahic, S., Mulak, K. E., & Escudero, P. (2017). Acoustic properties predict perception of unfamiliar Dutch vowels by adult Australian English and Peruvian Spanish listeners. *Frontiers in Psychology, 8*. doi: 10.3389/fpsyg.2017.00052

Antoniou, M., & Wong, P. (2015). Poor phonetic perceivers are affected by cognitive load when resolving talker variability. *The Journal of the Acoustical Society of America, 138*(2), 571-574.

Arvaniti, A. (1999a). Cypriot Greek. *Journal of the International Phonetic Association, 29*(2), 173-178.

Arvaniti, A. (1999b). Standard Modern Greek. *Journal of the International Phonetic Association, 29*(2), 167-172.

Arvaniti, A. (2001). Comparing the phonetics of single and geminate consonants in Cypriot and Standard Greek. *Proceedings of the 4th International Conference on Greek Linguistics, Thessaloniki,* 37-44.

Arvaniti, A. (2006). Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek. *San Diego Linguistic Papers, 2*, 1-24.

Arvaniti, A. (2007). Greek phonetics: The state of the art. *Journal of Greek Linguistics, 8*(1), 97-208.

Arvaniti, A. (2010). A (brief) overview of the phonetics and phonology of Cypriot Greek. In A. Voskos, D. Goutsos & A. Mozer (Eds.), *The Greek language in Cyprus from antiquity to today* (pp. 107-124). Athens, Greece: University of Athens.

Baese-Berk, M. M. (2019). Interactions between speech perception and production during learning of novel phonemic categories. *Attention, Perception, & Psychophysics*, *81*, 981-1005.

Baese-Berk, M. M., & Samuel, A. G. (2016). Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory and Language*, *89*, 23-36.

Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non-native speech. *Language and Linguistics Compass, 14*(7), e12375.

Baker, W., & Trofimovich, P. (2001). Does perception precede production? Evidence from Korean-English bilinguals. *LACUS Forum, 27,* 273-284.

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*(1), 1-27.

Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics in Language Teaching*, *44*(3), 231-250.

Balas, A. (2018). English vowel perception by polish advanced learners of English. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique, 63*(3), 309-338.

Baltazani, M. (2007). Prosodic rhythm and the status of vowel reduction in Greek. *Selected Papers on Theoretical and Applied linguistics, 17*(1), 31-43.

Baquiran, C. L. C., & Nicoladis, E. (2020). A doctor's foreign accent affects perceptions of competence. *Health Communication, 35*(6), 726-730.

Baran-Łucarz, M. (2011). The relationship between language anxiety and the actual and perceived levels of foreign language pronunciation. *Studies in Second Language Learning and Teaching, 1*(4), 491-514.

Barriuso, T. A., & Hayes-Harb, R. (2018). High variability phonetic training as a bridge from research to practice. *CATESOL Journal, 30*(1), 177-194.

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy, 9*(1), 95-114.

Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1835-1842.

Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics, 36*(1), 1-6.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bent, T., Bradlow, A. R., & Smith, B. L. (2007). Segmental errors in different word positions and their effects on intelligibility of non-native speech. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 331-347). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bernhardt, B. M., Bacsfalvi, P., Adler-Bock, M., Modha, G, & Purves, B. (2013). The investigation of speech production: Experimental and quasi-experimental approaches. In N. Müller & M. Ball (Eds.), *Research methods in clinical linguistics and phonetics: A practical guide* (pp. 48-62). Hoboken: John Wiley & Sons.

Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech, 46*(2-3), 183-216.

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

Best, C. T., & Tyler, M. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam/Philadelphia: John Benjamins Publishing Company.

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

Bettoni-Techio, M., Rauber, A. S., & Koerich, R. D. (2007). Perception and production of word-final alveolar stops by Brazilian Portuguese learners of English. *Proceedings of Interspeech 2007, Belgium,* 2293-2296.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 99-116). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44*(2), 235-249.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 6.1.37) [Software]. Available from http://www.praat.org/

Bogacka, A. (2004). On the perception of English high vowels by Polish learners of English. *CamLing 2004: Proceedings of the University of Cambridge Second Postgraduate Conference in Language Research, Cambridge,* 43-50.

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

Bohn, O. S., & Bundgaard-Nielsen, R. L. (2008). Second language speech learning with diverse inputs. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 207-218). Bristol: Multilingual Matters.

Bohn, O. S., & Flege, J. (1997). Perception and production of a new vowel category by adult second language learners. In A. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 53-74). Berlin, New York: Mouton de Gruyter.

Bohn, O., & Steinlen, A. K. (2003). Consonantal context affects cross-language perception of vowels. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona,* 2289-2292.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133-159). Mahwah: Erlbaum.

Bongaerts, T., Mennen, S., & Slik, F. v. d. (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*(2), 298-308.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30-50). Clevedon: Multilingual Matters.

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*(4), 447-465.

Botinis, A., Fourakis, M., & Prinou, I. (2000). Acoustic structure of the Greek stop consonants. *Glossologia*, *11-12*, 167-199.

Bozorgian, H., & Fallahpour, S. (2015). Teachers' and students' amount and purpose of L1 use: English as foreign language (EFL) classrooms in Iran. *Iranian Journal of Language Teaching Research, 3*(2), 67-81.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify english /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*(5), 977-985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV: Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*(4), 2299-2310.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*(1), 51-61.

Brekelmans, G., Lavan, N., Saito, H., Clayards, M., & Wonnacott, E. (2022). Does high variability training improve the learning of non-native phoneme contrasts over low variability training? A replication. *Journal of Memory and Language, 126*, 104352. doi: 10.1016/j.jml.2022.104352

Brosseau-Lapré, F., Rvachew, S., Clayards, M., & Dickson, D. (2013). Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics*, *34*(3), 419-441.

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly, 22*(4), 593-606.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1-19.

Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020). Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Quarterly Journal of Experimental Psychology, 73*(9), 1431-1443.

Buckingham, L. (2014). Attitudes to English teachers' accents in the gulf. *International Journal of Applied Linguistics, 24*(1), 50-73.

Cargile, A. C. (2000). Evaluations of employment suitability: Does accent always matter? *Journal of Employment Counseling, 37*(3), 165-177.

Carlet, A. (2017). *L2 perception and production of English consonants and vowels by Catalan speakers: The effects of attention and training task in a cross-training study* (Doctoral dissertation). Universitat Autonoma de Barcelona. Retrieved from https://www.tdx.cat/bitstream/handle/10803/403758/afc1de1.pdf?sequence=1

Carlet, A. (2019). Different high variability procedures for training L2 vowels and consonants. *Proceedings of the 19th International Congress of Phonetic Sciences, Canberra, Australia,* 944-948.

Carlet, A., & Cebrian, J. (2014). Training Catalan speakers to identify L2 consonants and vowels: A short-term high variability training study. *Proceedings of the 7th International Symposium on the Acquisition of Second Language Speech, 5,* 85-98.

Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling, 43*(2), 70-83.

Casillas, J. V. (2019). Phonetic category formation is perceptually driven during the early stages of adult L2 development. *Language and Speech*, *63*(3), 550-581.

Cebrian, J. (2006). Experience and the use of duration in the categorization of L2 vowels. *Journal of Phonetics, 34*(3), 372-387.

Cebrian, J. (2007). Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 3,* 1637-1640.

Cebrian, J. (2009). Effects of native language and amount of experience on crosslinguistic perception. *The Journal of the Acoustical Society of America*, *125*(4), 2775-2775.

Cebrian, J. (2019). Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. *The Journal of the Acoustical Society of America, 145*(1), EL52-EL58.

Cebrian, J., Carlet, A., Gorba, A., & Gavaldà, N. (2019). Perceptual training affects L2 perception but not cross-linguistic similarity. *Proceedings of the 19th International Congress of Phonetic Sciences, Canberra, Australia*, 929-933.

Cenoz, J., & Lecumberri, L. G. (1999). The effect of training on the discrimination of English vowels. *International Review of Applied Linguistics in Language Teaching, 37*(4), 261-276.

Chang, C. B., & Bowles, A. R. (2015). Context effects on second-language learning of tonal contrasts. *The Journal of the Acoustical Society of America, 138*(6), 3703-3716.

Cheng, B., Zhang, X., Fan, S., & Zhang, Y. (2019). The role of temporal acoustic exaggeration in high variability phonetic training: A behavioral and ERP study. *Frontiers in Psychology, 10*. doi: 10.3389/fpsyg.2019.01178

Clopper, C. G. (2009). Computational methods for normalizing acoustic vowel data for talker differences. *Language and Linguistics Compass*, *3*(6), 1430-1442.

Coutsougera, P. (2007). The impact of orthography on the acquisition of L2 phonology: Inferring the wrong phonology from print. *Proceedings of the Phonetics Teaching and*

Learning Conference, London, UK. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c73c7079150fb46f2bb3f51c595b345b643c1c3b

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). London: Sage Publications.

Crosby, C. (2020). *Adding production to high variability phonetic training* (Honors thesis). Sally McDonnell Barksdale Honors College, The University of Mississippi, Oxford. Retrieved from https://egrove.olemiss.edu/cgi/viewcontent.cgi?article=2478&context=hon_thesis

Čubrović, B., & Bjelaković, A. (2020). Pronunciation model selection, or do you speak American. *BELLS90 Proceedings: International Conference to Mark the 90th Anniversary of the English Department, Faculty of Philology, University of Belgrade, 1,* 139-151.

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, *111*(6), 2862-2873.

Dauer, R. M. (1980). The reduction of unstressed high vowels in Modern Greek. *Journal of the International Phonetic Association, 10*(1-2), 17-27.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499-533.

DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88-108). New York: Oxford University Press.

Derwing, T. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59*(4), 547-567.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review, 66*(2), 181-202.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30*(2), 155-166.

Derwing, T. M., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning, 64*(3), 526-548.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, *54*(4), 655-679.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association, 27*(1-2), 47-55.

Detey, S., & Racine, I. (2015). Does perception precede production in the initial stage of French nasal vowel quality acquisition by Japanese learners? A corpus-based discrimination experiment. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.

Detey, S., Racine, I., Eychenne, J., & Kawaguchi, Y. (2014). Corpus-based L2 phonological data and semi-automatic perceptual analysis: the case of nasal vowels produced by beginner Japanese learners of French. *Proceedings of Interspeech 2014, Singapore,* 539-543.

Di Paolo, M., Yaeger-Dror, M., & Wassink, A. B. (2011). Analyzing vowels. In M. Di Paolo & M. Yaeger-Dror (Eds.), *Sociophonetics: A student's guide* (pp. 87-106). London: Routledge.

Dimitriou, D. (2019). L2 acquisition and production of the English rhotic by L1 Greek-Cypriot speakers: The effect of L1 articulatory routines and phonetic context. *Philologia, 16*(16), 45-64.

Dimitriou, D. (2022). Perceptual training effects on the perception and production of L2 English vowels by L1 Greek-Cypriot learners. In E. Kkese (Ed.), *Speech perception and production in L2* (pp. 125-151). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris Publications.

Dong, H., Clayards, M., Brown, H., & Wonnacott, E. (2019). The effects of high versus low talker variability and individual aptitude on phonetic training of Mandarin lexical tones. *PeerJ*, *7*, e7191.

Dornbusch, T. (2012). *Orthographic influences on L2 auditory word processing* (Doctoral dissertation). Technische Universität Dortmund, Germany. Retrieved from https://eldorado.tu-dortmund.de/bitstream/2003/29641/1/Dissertation.pdf

Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, *100*, 19-47.

Dragojevic, M., & Goatley-Soan, S. (2020). Americans' attitudes toward foreign accents: Evaluative hierarchies and underlying processes. *Journal of Multilingual and Multicultural Development, 43*(2), 167-181.

Dragojevic, M., Giles, H., Beck, A., & Tatum, N. T. (2017). The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs, 84*(3), 385-405.

DuBois, S. (2018). Identification and stereotypes of foreign-accented English. *Notandum, 46*, 7-28.

Elliott, A. R. (1995). Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal, 79*(3), 356-371.

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191-228.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands: Netherlands Graduate School of Linguistics, LOT Publications.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(4), 551-585.

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America, 128*(5), EL254-EL260.

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics, 37*(4), 452-465.

Escudero, P., Benders, T., & Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second language vowels. *The Journal of the Acoustical Society of America*, *130*(4), EL206-EL212.

Evans, B. G., & Martín-Alvarez, L. (2016). Age-related differences in second-language learning? A comparison of high and low variability perceptual training for the acquisition of English /i/-/ɪ/ by Spanish adults and children. *Proceedings of New Sounds: 8th International Conference on Second Language Speech.* Retrieved from https://conferences.au.dk/newsounds2016

Fabra, L. R. (2005). Predicting ease of acquisition of L2 speech sounds. A perceived dissimilarity test. *Vigo International Journal of Applied Linguistics*, *2*, 75-92.

Fabra, L. R., & Romero, J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics, 40*(3), 491-508.

Ferguson, C. A. (1959). Diglossia. *Word, 15*(2), 325-340.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America, 93*(3), 1589-1608.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E. (1999). The relation between L2 production and perception. *Proceedings of the 14th International Congress of Phonetics Sciences, San Francisco, 2,* 1273-1276.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister (Ed.), *An integrated view of language development. Papers in honor of Henning Wode* (pp. 217-244). Trier: Wissenschaftlicher Verlag.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.), *Laboratory phonology 9* (pp. 353-381). Berlin: Mouton de Gruyter.

Flege, J. E. (2008). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175-190). Bristol: Multilingual Matters.

Flege, J. E. (2018). It's input that matters most, not age. *Bilingualism: Language and Cognition, 21*(5), 919-920.

Flege, J. E., & Bohn, O. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge: Cambridge University Press.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23*(4), 527-552.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*(1), 1-34.

Flege, J. E., & MacKay, I. R. A. (2011). What accounts for “age” effects on overall degree of foreign accent? In M. Wrembel, M. Kul & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New sounds* (pp. 65-82). Bern, Switzerland: Peter Lang.

Flege, J. E., & Wayland, R. (2019). The role of input in native Spanish late learners’ production and perception of English phonetic segments. *Journal of Second Language Studies, 2*(1), 1-44.

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*(2), 153-175.

Flege, J. E., Bohn, O., & Jang, S. (1997b). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997a). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*(2), 169-186.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999b). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106*(5), 2973-2987.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*(1), 1-26.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*(4), 467-491.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999a). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104.

Flynn, N. (2011). Comparing vowel formant normalisation procedures. *York Papers in Linguistics*, *2*(11), 1-28.

Fourakis, M. (1986). An acoustic study of the effects of tempo and stress on segmental intervals in Modern Greek. *Phonetica, 43*(4), 172-188.

Fourakis, M., Botinis, A., & Katsaiti, M. (1999). Acoustic characteristics of Greek vowels. *Phonetica, 56*(1-2), 28-43.

Garcia Perez, G. M. (2003). *Training Spanish speakers in the perception and production of English vowels* (Doctoral dissertation). Simon Fraser University, Canada. Retrieved from https://summit.sfu.ca/_flysystem/fedora/sfu_migrate/8727/b30941805.pdf

Gardner, R. C., & Lambert, W. E. (1972). Attitudes and motivation in second-language learning. Rowley, MA: Newbury House.

George, U., & Chaze, F. (2014). Discrimination at work: Comparing the experiences of foreign-trained and locally-trained engineers in Canada. *Canadian Ethnic Studies, 46*(1), 1-21.

Georgiou, G. P. (2019). Bit and beat are heard as the same: Mapping the vowel perceptual patterns of Greek-English bilingual children. *Language Sciences, 72*, 1-12.

Georgiou, G. P. (2021). The impact of auditory perceptual training on the perception and production of English vowels by Cypriot Greek children and adults. *Language Learning and Development, 18*(4), 379-392.

Giannakopoulou A., Uther M., & Ylinen S. (2013a). Enhanced plasticity in spoken language acquisition for child learners: evidence from phonetic training studies in child and adult learners of English. *Child Language Teaching and Therapy*, *29*(2), 201-218.

Giannakopoulou A., Uther M., & Ylinen S. (2013b). Phonetic and orthographic cues are weighted in speech sound perception by second language speakers: evidence from Greek speakers of English. *The Journal of the Acoustical Society of America*, *134*(5), 1-9.

Giannakopoulou, A., Brown, H., Clayards, M., & Wonnacott, E. (2017). High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ, 5*, e3209.

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science, 2*(3), 119-128.

Gluszek, A., & Dovidio, J. F. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology, 29*(2), 224-234.

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311-343.

Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *Journal of Phonetics, 28*(1), 27-42.

Hammerly, H. (1982). Contrastive phonology and error analysis. *International Review of Applied Linguistics, 20,* 17-32.

Hao, Y., & de Jong, K. (2016). Imitation of second language sounds in relation to L2 perception and production. *Journal of Phonetics, 54*, 151-168.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495-522.

Hattori, K., & Iverson, P. (2009). English/r/-/l/category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *The Journal of the Acoustical Society of America, 125*(1), 469-479.

Hattori, K., & Iverson, P. (2010). Examination of the relationship between L2 perception and production: An investigation of English/r/-/l/perception and production by adult Japanese speakers. *Second Language Studies: Acquisition, Learning, Education and Technology, Tokyo, Japan*. Retrieved from https://www.isca-speech.org/archive_v0/L2WS_2010/papers/lw10_P2-4.pdf

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*(4), 458-484.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech, 53*(3), 367-381.

Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., Iba, M., & Chung, H. (2006). The use of visual cues in the perception of non-native consonant contrasts. *The Journal of the Acoustical Society of America, 119*(3), 1740-1751.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication, 47*(3), 360-378.

Hendriks, B., van Meurs, F., & Hogervorst, N. (2016). Effects of degree of accentedness in lecturers' Dutch-English pronunciation on Dutch students' attitudes and perceptions of comprehensibility. *Dutch Journal of Applied Linguistics, 5*(1), 1-17.

Hendriks, B., van Meurs, F., & Reimer, A. (2018). The evaluation of lecturers' nonnative-accented English: Dutch and German students' evaluations of different degrees of

Dutch-accented and German-accented English of lecturers in higher education. *Journal of English for Academic Purposes, 34*, 28-45.

Hendriks, B., van Meurs, F., & Usmany, N. (2021). The effects of lecturers' non-native accent strength in English on intelligibility and attitudinal evaluations by native and non-native English students. *Language Teaching Research,* 1-30. doi: 10.1177/1362168820983145

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America, 108*(6), 3013-3022.

Hoffman, M. (2014). Sociolinguistic interviews. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 25-41). Hoboken: John Wiley & Sons.

Holliday, J. J. (2010). Inter-and intra-L1 differences in L2 speech perception. *Proceedings of Second Language Studies: Acquisition, Learning, Education and Technology, Tokyo, Japan.* Retrieved from https://www.isca-speech.org/archive_v0/L2WS_2010/papers/lw10_O1-2.pdf

Hosoda, M., Nguyen, L. T., & Stone-Romero, E. F. (2012). The effect of Hispanic accents on employment decisions. *Journal of Managerial Psychology, 27*(4), 347-364.

Hu, W., Mi, L., Yang, Z., Tao, S., Li, M., Wang, W., ... & Liu, C. (2016). Shifting perceptual weights in L2 vowel identification after training. *PloS One*, *11*(9), e0162876. doi: 10.1371/journal.pone.0162876

Hu, X., Kuo, L., & Dixon, L. Q. (2022). Effectiveness of interventions on improving English language learners' intelligibility and comprehensibility: A meta-analytic review. *Language Teaching Research*, 1-25. doi: 10.1177/13621688221106832

Huensch, A., & Tremblay, A. (2015). Effects of perceptual phonetic training on the perception and production of second language syllable structure. *Journal of Phonetics, 52*, 105-120.

Hughes, A., Trudgill, P., & Watt, D. (2013). *English accents and dialects: An introduction to social and regional varieties of English in the British Isles* (5th ed.). New York: Routledge.

Hutchinson, A. E., & Dmitrieva, O. (2022). Exposure to speech via foreign film and its effects on non-native vowel production and perception. *Journal of Phonetics, 95*, 101189. doi: 10.1016/j.wocn.2022.101189

Immonen, K., Peltola, K. U., Tamminen, H., Alku, P., & Peltola, M. S. (2023). Orthography does not hinder non-native production learning in children. *Second Language Research, 39*(2), 565-577.

Inceoglu, S. (2016). Effects of perceptual training on second language vowel perception and production. *Applied Psycholinguistics, 37*(5), 1175-1199.

Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*(2), 135-159.

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*(3), 475-505.

Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America, 122*(5), 2842-2854.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America, 126*(2), 866-877.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/to Japanese adults. *The Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57.

Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics, 33*(1), 145-160.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America, 119*(2), 1118-1130.

Jin, S. H., & Liu, C. (2014). Intelligibility of American English vowels and consonants spoken by international students in the United States. *Journal of Speech, Language, and Hearing Research*, *57*(2), 583-596.

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*(1), 60-99.

Jongman, A., Fourakis, M., & Sereno, J. A. (1989). The acoustic vowel space of Modern Greek and German. *Language and Speech, 32*(3), 221-248.

Jügler, J., Zimmerer, F., Möbius, B., & Draxler, C. (2015). The effect of high-variability training on the perception and production of French stops by German native speakers. *Proceedings of Interspeech 2015, Dresden, Germany,* 806-810.

Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01246

Kartushina, N., & Martin, C. D. (2019). Talker and acoustic variability in learning to produce nonnative sounds: Evidence from articulatory training. *Language Learning, 69*(1), 71-105.

Katamba, F. (2009). Segmental phonology. In J. Culpeper, F. Katamba, P. Kerswill, R. Wodak & T. McEnery (Eds.), *English language: Description, variation and context* (pp. 30-54). Basingstoke, Hampshire: Palgrave Macmillan.

Kelley, M. C., & Tucker, B. V. (2020). A comparison of four vowel overlap measures. *The Journal of the Acoustical Society of America*, *147*(1), 137-145.

Kerswill, P., & Watson, K. (2014). Phonological considerations in sociophonetics. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 136-148). Hoboken: John Wiley & Sons.

Kkese, E., & Petinou, K. (2017). Perception abilities of L1 Cypriot Greek listeners – Types of errors involving plosive consonants in L2 English. *Journal of Psycholinguistic Research, 46*, 1-25.

Kluge, D. C., Rauber, A. S., Reis, M. S., & Bion, R. A. H. (2007). The relationship between the perception and production of English nasal codas by Brazilian learners of English. *Proceedings of Interspeech 2007, Belgium,* 2297-2300.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research, 18*, 201-222.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics, 38*(4), 569-587.

Koutsoudas, A., & Koutsoudas, O. (1962). A contrastive analysis of the segmental phonemes of Greek and English. *Language Learning, 12*(3), 211-230.

Krashen, S. D. (1975). The critical period for language acquisition and its possible bases. *Annals of the New York Academy of Sciences, 263*(1), 211-224.

Krzonowski, J., Ferragne, E., & Pellegrino, F. (2015). Effects of perception and production trainings on the production of English vowels by French native learners. In A. Botinis (Ed.), *Proceedings of the 6th tutorial and research workshop on experimental linguistics* (pp. 37-40). Athens, Greece: National and Kapodistrian University of Athens.

Kyprianou, M. (2007). Teaching pronunciation in Cyprus. *Proceedings of the Phonetics Teaching and Learning Conference, London, UK.* Retrieved from https://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_39e.pdf

Kyprianou, M. (2015). *Teaching and learning English pronunciation in Cyprus: Survey, analysis and challenges* (Doctoral dissertation). University of Cyprus.

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed.). USA: Wadsworth.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American

English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*(2), 227-247.

Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition*, *13*(3)*,* 305-341.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345-366.

Lehn, W., & Slager, W. R. (1959). A contrastive study of Egyptian Arabic and American English: The segmental phonemes. *Language Learning, 9*(1-2), 25-33.

Lengeris, A. (2008). The effectiveness of auditory phonetic training on Greek native speakers' perception and production of southern British English vowels. In A. Botinis (Ed.), *ExLing 2008: Proceedings of the 2nd Tutorial and Research Workshop on Experimental Linguistics* (pp. 133-136). Athens, Greece: National and Kapodistrian University of Athens.

Lengeris, A. (2009a). *Individual differences in second-language vowel learning* (Doctoral dissertation). University College London, UK. Retrieved from https://discovery.ucl.ac.uk/id/eprint/19029/1/19029.pdf

Lengeris, A. (2009b). Perceptual assimilation and L2 learning: Evidence from the perception of southern British English vowels by native speakers of Greek and Japanese. *Phonetica, 66*(3), 169-187.

Lengeris, A. (2018). Computer-based auditory training improves second-language vowel production in spontaneous speech. *The Journal of the Acoustical Society of America, 144*(3), EL165-EL171.

Lengeris, A., & Hazan, V. (2007). Cross-language perceptual assimilation and discrimination of southern British English vowels by Greek and Japanese learners of English. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 3,* 1641-1644.

Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *The Journal of the Acoustical Society of America, 128*(6), 3757-3768.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley and Sons.

Leung, K. K., Jongman, A., Wang, Y., & Sereno, J. A. (2016). Acoustic characteristics of clearly spoken English tense and lax vowels. *The Journal of the Acoustical Society of America*, *140*(1), 45-58.

Levy, E. S., & Law, F. F. (2010). Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. *The Journal of the Acoustical Society of America, 128*(3), 1290-1305.

Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, *36*(1), 141-157.

Lipińska, D. (2017). Production of English /i:/and/ɪ/ by Polish learners of English: A longitudinal study. In E. Babatsouli (Ed.), *Proceedings of the international symposium on monolingual and bilingual speech* (pp. 180-185). Chania, Greece: Institute of Monolingual and Bilingual Speech.

Lippi-Green, R. (2012). *English with an accent: Language, ideology and discrimination in the United States* (2nd ed.). London: Routledge.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242-1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III: Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America, 96*(4), 2076-2087.

Logan, J. S., & Pruitt, J. S. (1995). Methodological issues in training listeners to perceive non-native phonemes. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 351-377). Timonium, MD: York Press.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874-886.

Long, M. (2005). Problems with supposed counter-evidence to the critical period hypothesis. *International Review of Applied Linguistics in Language Teaching, 43*(4), 287-317.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*(3), 251-285.

Lopez-Soto, T., & Kewley-Port, D. (2009). Relation of perception training to production of codas in English as a second language. *Proceedings of Meetings on Acoustics, 6*(1), 062003. doi: 10.1121/1.3262006

Loukina, A. (2009). Phonetic variation in spontaneous speech: Vowel and consonant reduction in Modern Greek dialects. In O. P. Jones & E. Payne (Eds.), *Oxford university working papers in linguistics, philology & phonetics: Papers in phonetics and computational linguistics* (pp. 36-56). University of Oxford: Citeseer.

Mackridge, P. A. (1985). *The Modern Greek language: A descriptive analysis of standard Modern Greek*. Oxford: Oxford University Press.

Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, *13*(3), 279-297.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*(2), 229-258.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(2), 89-108.

McEntee-Atalianis, L. J. (2004). The impact of English in postcolonial, postmodern Cyprus. *International Journal of the Sociology of Language, 168*, 77-90.

McEntee-Atalianis, L. J., & Pouloukas, S. (2001). Issues of identity and power in a Greek-Cypriot community. *Journal of Multilingual and Multicultural Development, 22*(1), 19-38.

Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from anglophones' processing of the french/u/–/y/contrast. *Second Language Research, 38*(3), 581-605.

Mildner, V. (2013). Experimental and quasi-experimental research in clinical linguistics and phonetics. In N. Müller & M. Ball (Eds.), *Research methods in clinical linguistics and phonetics: A practical guide* (pp. 28-47). Hoboken: John Wiley & Sons.

Mligo, E. S. (2016). *Introduction to research methods and report writing: A practical guide for students and researchers in social sciences and the humanities*. Eugene, Oregon: Resource Publications.

Mora, J. C., & Fullana, N. (2007). Production and perception of English /i:/-/ɪ/ and /æ/-/ʌ/ in a formal setting: Investigating the effects of experience and starting age. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 3,* 1613-1616.

Morrison, G. S. (2002). *Effects of L1 duration experience on Japanese and Spanish listeners' perception of English high front vowels* (Master's thesis). Simon Fraser University, Canada. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2f7be7f64eec2bed8b027fddd8c787e340dd5e35

Morrison, G. S. (2003). Perception and production of Spanish vowels by English speakers. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona,* 1533-1536.

Motohashi-Siago, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology, 13*(2), 29-47.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*(1), 81-108.

Moyer, A. (2004). *Age, accent, and experience in second language acquisition: An integrated approach to critical period inquiry*. 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), 502-518.

Moyer, A. (2008). Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 159-174). Bristol: Multilingual Matters.

Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge: Cambridge University Press.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*(2), 38-51.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*(4), 520-531.

Munro, M. J., Derwing, T. M., & Burgess, C. S. (2010). Detection of nonnative speaker status from content-masked speech. *Speech Communication*, *52*(7-8), 626-637.

Munro, M. J., Derwing, T. M., & Thomson, R. I. (2015). Setting segmental priorities for English learners: Evidence from a longitudinal study. *International Review of Applied Linguistics in Language Teaching, 53*(1), 39-60.

Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System, 42*, 258-269.

Nagle, C. (2018, August 31). Perception, production, and perception-production: Research findings and implications for language pedagogy. *Contact Magazine*, pp. 5-12. Retrieved from http://contact.teslontario.org/perception-production/

Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition, 44*(2), 580-605.

Neri, A., Cucchiarini, C., & Strik, H. (2006). Selecting segmental errors in non-native Dutch for optimal pronunciation training. *International Review of Applied Linguistics in Language Teaching, 44*(4), 357-404.

Nicolaidis, K. (2003). Acoustic variability of vowels in Greek spontaneous speech. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona,* 3221-3224.

Nimz, K., & Khattab, G. (2020). On the role of orthography in L2 vowel production: The case of polish learners of German. *Second Language Research, 36*(4), 623-652.

Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research, 50*(6), 1496-1509.

Nishi, K., & Kewley-Port, D. (2008). Nonnative speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research, 51*(6), 1480-1493.

Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America*, *124*(1), 576-588.

Okuno, T., & Hardison, D. M. (2016). Perception-production link in L2 Japanese vowel duration: Training with technology. *Language Learning & Technology, 20*(2), 61-80.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*(3), 261-283.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*, 533-544.

Pan, Y., & Pan, Y. (2010). The use of L1 in the foreign language classroom. *Colombian Applied Linguistics Journal, 12*(2), 87-96.

Papachristou, V. (2011). Explicit vs. implicit pronunciation teaching to Greek children: The case of the acquisition of English vowels. *Selected Papers on Theoretical and Applied Linguistics, 19*, 371-381.

Patkowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics, 11*(1), 73-89.

Payne, A. (1980). Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In W. Labov (Ed.), *Locating language in time and space* (pp. 143-178). New York: Academic Press.

Peltola, K. U., Tamminen, H., Alku, P., & Peltola, M. S. (2015). Non-native production training with an acoustic model and orthographic or transcription cues. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.

Peperkamp, S., & Bouchon, C. (2011). The relation between perception and production in L2 phonological processing. *Proceedings of Interspeech 2011, Florence, Italy*, 161-164.

Perrachione, T. K., Lee, J., Ha, L. Y., & Wong, P. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America, 130*(1), 461-472.

Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 301-314). Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Piske, T., & Young-Scholten, M. (2008). *Input matters in SLA*. Bristol: Multilingual Matters.

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*(2), 191-215.

Pruitt, J. S., Jenkins, J. J., & Strange, W. (2006). Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *The Journal of the Acoustical Society of America, 119*(3), 1684-1696.

Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30*(2), 271-287.

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology, 22*(1), 69-96.

QSR International Pty Ltd (1999-2023). NVivo (Version 14) [Software]. Available from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

R Core Team (2022). R: A language and environment for statistical computing (Version 4.2.1) [Software]. Available from https://www.R-project.org/

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1588-1594.

Rato, A., & Carlet, A. (2020). Second language perception of English vowels by Portuguese learners: The effect of stimulus type. *Ilha do Desterro, 73*(3), 205-226.

Rato, A., & Rauber, A. S. (2015). The effects of perceptual training on the production of English vowel contrasts by Portuguese learners. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.

Rauber, A., Rato, A., Kluge, D., & Santos, G. (2012). TP – Perception Tests / Perception Training Tasks (Version 3.1) [Software]. Available from http://www.worken.com.br/tp/tp_install.html

Roach, P. (2004). British English: Received pronunciation. *Journal of the International Phonetic Association, 34*(2), 239-245.

Roessel, J., Schoel, C., Zimmermann, R., & Stahlberg, D. (2019). Shedding new light on the evaluation of accented speakers: Basic mechanisms behind nonnative listeners' evaluations of nonnative accented job candidates. *Journal of Language and Social Psychology, 38*(1), 3-32.

Rojczyk, A. (2010). Forming new vowel categories in second language speech: The case of Polish learners' production of English /ɪ/ and /e/. *Research in Language*, *8*, 85-97.

Sadakata, M., & McQueen, J. M. (2013). High stimulus variability in nonnative speech learning supports formation of abstract categories: Evidence from Japanese geminates. *The Journal of the Acoustical Society of America, 134*(2), 1324-1335.

Sadakata, M., & McQueen, J. M. (2014). Individual aptitude in mandarin lexical tone perception predicts effectiveness of high-variability training. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01318

Said, S. B. (2006). *Attitudes towards accented speech: A comparative study of native and non-native speakers of American English* (Master's thesis). Southern Illinois University,

Carbondale. Retrieved from https://www.proquest.com/openview/bc02ecc4e9813a857d579a1396026c3d/1?pq-origsite=gscholar&cbl=18750&diss=y

Saito, K., Kachlicka, M., Suzukida, Y., Petrova, K., Lee, B. J., & Tierney, A. (2022). Auditory precision hypothesis-L2: Dimension-specific relationships between auditory processing and second language segmental learning. *Cognition*, *229*, 105236. doi: 10.1016/j.cognition.2022.105236.

Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *38*(4), 439-462.

Sakai, M. (2016). *(Dis)connecting perception and production: Training native speakers of spanish on the english /iː/-/ɪ/ distinction* (Doctoral dissertation). Georgetown University, Washington. Retrieved from https://www.proquest.com/openview/d0693cd0c3bfdb572a2e3071ad259c38/1?pq-origsite=gscholar&cbl=18750

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics, 39*(1), 187-224.

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, *23*(3), 246-255.

Schleef, E. (2014). Written surveys and questionnaires in sociolinguistics. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics: A practical guide* (pp. 42-57). Hoboken: John Wiley & Sons.

Schmidt, R. (2001). Attention. In P. Robinson. (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Schouten, A. (2009). The critical period hypothesis: Support, challenge, and reconceptualization. *Studies in Applied Linguistics and TESOL, 9*(1), 1-16.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19*(3-4), 245-253.

Scovel, T. (2011). Age, acquisition, and accent. In C. Abello-Contesse, R. Chacon-Beltran & M. D. Lopez-Jimenez (Eds.), *Age in L2 acquisition and teaching* (pp. 31-48). Bern, Switzerland: Peter Lang.

Seliger, H. W. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 11-19). New York: Academic Press.

Sfakianaki, A. (2002). Acoustic characteristics of Greek vowels produced by adults and children. *Selected Papers on Theoretical and Applied linguistics*, *14*, 383-394.

Shinohara, Y., & Iverson, P. (2015). Effects of English/r/-/l/perceptual training on Japanese children's production. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow.*

Simaki, V., Mporas, I., & Kondyli, M. (2015). Acoustic differences between Standard Greek and Cypriot Greek. *Proceeding of the 17th International Conference on Speech and Computer, University of Patras,* 71-78.

Simon, E., Chambless, D., & Alves, U. K. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences, 32*(3), 380-394.

Singleton, D. (2003). Critical period or general age factor(s)? In M. P. G. Mayo & M. L. G. Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 3-22). Clevedon: Multilingual Matters.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching, 43*(4), 269-285.

Stockwell, R., & Bowen, J. (1965). *The sounds of English and Spanish*. Chicago: University of Chicago Press.

Stoehr, A., & Martin, C. D. (2022). The impact of orthographic forms on speech production and perception: An artificial vowel-learning study. *Journal of Phonetics, 94*, 101180. doi: 10.1016/j.wocn.2022.101180

Strange, W. (2007). Cross-language phonetic similarity of vowels: Theoretical and methodological issues. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 35-55). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 26*(2), 233-253.

Suzukida, Y., & Saito, K. (2019). Which segmental features matter for successful L2 comprehensibility? revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research, 25*(3), 431-450.

Szyszka, M. (2011). Foreign language anxiety and self-perceived English pronunciation competence. *Studies in Second Language Learning and Teaching, 1*(2), 283-300.

Themistocleous, C. (2017a). Dialect classification using vowel acoustic parameters. *Speech Communication, 92*, 13-22.

Themistocleous, C. (2017b). The nature of phonetic gradience across a dialect continuum: Evidence from Modern Greek vowels. *Phonetica, 74*(3), 157-172.

Themistocleous, C., & Logotheti, A. (2016). Standard Modern Greek and Cypriot Greek vowels: A sociophonetic study. *Modern Greek Dialects and Linguistics Theory, 6*(1), 178-184.

Thomas, E. R., & Kendall, T. (2015a). NORM: The vowel normalisation and plotting suite (Version 1.1) [Web-based interface]. Available from http://lingtools.uoregon.edu/norm/norm1.php

Thomas, E. R., & Kendall, T. (2015b). *NORM's vowel normalization Methods: Lobanov Method(s)*. Retrieved August 18, 2023, from http://lingtools.uoregon.edu/norm/norm1_methods.php#lobanov

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*(2), 177-204.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal, 28*(3), 744-765.

Thomson, R. I. (2012). Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Language Learning, 62*(4), 1231-1258.

Thomson, R. I. (2018). High variability [pronunciation] training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation, 4*(2), 208-231.

Thomson, R. I., & Derwing, T. M. (2016). Is phonemic training using nonsense or real words more effective? In J. Levis, H. Le, I. Lucic, E. Simpson & S. Vo (Eds.), *Proceedings of the 7th pronunciation in second language learning and teaching conference* (pp. 88-97). Ames, IA: Iowa State University.

Thomson, R. I., & Isaacs, T. (2009). Within-category variation in L2 English vowel learning. *Canadian Acoustics, 37*(3), 138-139.

Tsang, A. (2022). The relationships between EFL learners' anxiety in oral presentations, self-perceived pronunciation, and speaking proficiency. *Language Teaching Research.* doi: 10.1177/13621688221102522

Tyler, M. D. (2019). PAM-L2 and phonological category acquisition in the foreign language classroom. In A. M. Nyvad, M. Hejná, A. Højen, A. B. Jespersen & M. H. Sørensen (Eds.), *A sound approach to language matters – In honor of Ocke-Schwen Bohn* (pp. 607-630). Denmark: Aarhus University.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica, 71*(1), 4-21.

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience, 8*. doi: 10.3389/fnhum.2014.00577

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica, 64*(2-3), 122-144.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*(4), 539-552.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*(6), 3649-3658.

Watson, K. (2009). Regional variation in English accents and dialects. In J. Culpeper, F. Katamba, P. Kerswill, R. Wodak & T. McEnery (Eds.), *English language: Description, variation and context* (pp. 337-357). Basingstoke, Hampshire: Palgrave Macmillan.

Wells, J. C. (2005). Goals in teaching English pronunciation. In K. Dziubalska-Kołaczyk & J. Przedlacka (Eds.), *English pronunciation models: A changing scene* (pp. 101-110). Bern, Switzerland: Peter Lang.

Wong, J. W. S. (2012). Training the perception and production of English /e/ and /æ/ of Cantonese ESL learners: A comparison of low vs. high variability phonetic training. *Proceedings of the 14th Australasian International Conference on Speech Science and Technology (SST 2012), Sydney, Australia,* 37-40.

Wong, J. W. S. (2013). The effects of perceptual and/or productive training on the perception and production of English vowels /ɪ/ and /i:/ by Cantonese ESL learners. *Proceedings of Interspeech 2013, Lyon, France,* 2113–2117.

Wong, J. W. S. (2014). The effects of high and low variability phonetic training on the perception and production of English vowels /e/-/æ/ by Cantonese ESL learners with high and low L2 proficiency levels. *Proceedings of Interspeech 2014, Singapore,* 524-528.

Wong, J. W. S. (2015). Comparing the perceptual training effects on the perception and production of English high-front and high-back vowel contrasts by Cantonese ESL learners. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.

Wright, S. (1996). Accents of English. In D. Graddol, D. Leith & J. Swann (Eds.), *English: History, diversity and change* (pp. 259-300). London: Routledge.

Yazawa, K., Whang, J., Kondo, M., & Escudero, P. (2020). Language-dependent cue weighting: An investigation of perception modes in L2 learning. *Second Language Research, 36*(4), 557-581.

Yazgin, N. (2007). *The role of the English language in Cyprus and its effects on the ELT classroom* (Online submission). *Retrieved from* ERIC Digest (https://files.eric.ed.gov/fulltext/ED496971.pdf)

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience, 22*(6), 1319-1332.

Yousofi, N., & Naderifarjad, Z. (2015). The relationship between motivation and pronunciation: A case of Iranian EFL learners. *Journal of Applied Linguistics and Language Research, 2*(4), 249-262.

Yuan, Q., & Archibald, J. (2022). Modified input training and cue reweighting in second language vowel perception. *Frontiers in Educational Research, 5*(6), 65-75.

Zhang, K., & Peng, G. (2017). The relationship between the perception and production of non-native tones. *Proceedings of Interspeech 2017, Stockholm, Sweden,* 1799-1803.

Zhang, X., Cheng, B., Qin, D., & Zhang, Y. (2021). Is talker variability a critical component of effective phonetic training for nonnative speech? *Journal of Phonetics, 87*, 101071. doi: 10.1016/j.wocn.2021.101071

Zhi, N., & Li, A. (2021). Acoustic salience in the evaluations of intelligibility and foreign accentedness of nonnative vowel production. *Lingua, 256*, 103069. doi: 10.1016/j.lingua.2021.103069

Appendix A

Code	Gender	Age	Degree title	AOL	Years of learning English
CYG01	Male	20	Web Design and Development	7	10
CYG03	Female	18	Business Administration (Accounting and Finance)	11	7
CYG04	Female	20	Accounting and Finance	8	10
CYG05	Female	19	Business Administration	8	9
CYG06	Male	18	Accounting and Finance	8	9
CYG07	Male	24	Business Administration	12	8
CYG08	Female	25	Psychology	8	8
CYG09	Male	20	Hospitality and Tourism Management	9	7
CYG10	Female	23	Medicine	9	9
CYG11	Female	18	Accounting and Finance	12	6
CYG12	Female	18	English Language Studies	7	11
CYG15	Male	24	Business Computing	8	8
CYG16	Male	20	Accounting and Finance	9	10
CYG17	Male	20	Computer Science	8	8

Table A1. CYG participants' information

Code	Gender	Age	Degree title
NE01	Female	26	Psychology
NE02	Female	24	Psychology
NE03	Female	25	Sociology
NE04	Male	25	Linguistics
NE05	Female	28	Linguistics
NE06	Female	19	TESOL
NE07	Female	21	English Language and Linguistics
NE08	Male	20	Electrical Engineering
NE09	Male	22	Accounting and Finance
NE10	Male	24	Business Administration

Table A2. NE participants' information

Appendix B

B.1 Questionnaire completed by CYG participants

Linguistic Background Questionnaire

Please fill in this questionnaire in full and as honestly as possible. Please note that you may ask the researcher for clarifications if you feel that a question is confusing.

It should not take more than 20 minutes to complete the questionnaire. Thank you in advance for taking the time to complete this questionnaire.

1. Name (Code): _____

2. Gender: □ Male □ Female □ Prefer not to say

3. Age: _____

4. Year of Studies:

□ 1st □ 2nd □ 3rd □ 4th □ PG student □ Graduate □ Other:_____

Month:_____

5. University Background: □ English-speaking □ (Cypriot-)Greek-speaking

6. Degree Title: _____

7. Primary School Background: □ English-speaking □ (Cypriot-)Greek-speaking

8. Secondary School Background: □ English-speaking □ (Cypriot-)Greek-speaking

9. Age you started learning English: _____

10. Do you have any relatives who live in English-speaking countries? □ Yes □ No

11. If yes, in which countries do they live? _____
_____

12. How often do you see your relatives who live in English-speaking countries? Please state approximately a) how often they visit, b) for how long, c) how many days you spend with them (e.g. in a year) and d) what languages you speak when you are together (if applicable):_____
_____
_____

13. Approximately how often do you visit an English-speaking country? Please state the countries and frequency of visits: _____
_____
_____

14. Approximately how long do you stay during these visits?
_____
_____

15. If you left Cyprus and lived elsewhere for a period of more than 1 month, please state the place you lived in and for how long:

_____
_____
_____

16. For how many years have you been learning English? _____

17. Where and how have you learnt English? (Select all that apply)

□ (Cypriot-)Greek-speaking school
□ English-speaking school
□ Afternoon lessons (tutor centres, institutes)
□ At home
□ With friends
□ With relatives
□ Through movies/series
□ Other (please state): _____

18. Please provide information about any English language certificates that you have obtained (e.g. IGCSE, IELTS, entry to university exams etc.):

Certificate(s): _____**Year(s):**_____

Grades (please include speaking, reading, writing, listening grades if available): _____
_____

19. Please rate your English language skills:

	Poor	Below Average	Average	Above Average	Very Good	Excellent
Understanding						
Speaking						
Writing						
Listening						

20. (Cypriot-)Greek is my native language. □ Yes □ No

21. (Cypriot-)Greek is my parents' native language. □ Yes □ No

22. I was born and brought up in a (Cypriot-)Greek community where the main language spoken is (Cypriot-)Greek. □ Yes □ No

23. Please rate the following statements (1-never, 6-all the time). During my school years,

I used my native language at home. 1 2 3 4 5 6
I used my native language in social settings. 1 2 3 4 5 6
I used other languages at home. 1 2 3 4 5 6
I used other languages in social settings. 1 2 3 4 5 6

24. If you used any languages other than (Cypriot-)Greek at home or in social settings during your school years, please explain (what languages, why, how often, with whom):

_____
_____
_____

25. When you were at <u>middle school and high school,</u> how often did you use (spoke or listened to) <u>(Cypriot-)Greek</u>

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

26. During your <u>university years</u>, how often do you use (speak or listen to) <u>(Cypriot-)Greek</u>

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

27. When you were at <u>middle school and high school,</u> how often did you use (spoke or listened to) <u>English</u>

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in the English classroom?						
b) in other classes?						
c) at home?						
d) in social settings?						

28. How much of your use of English during your school years occurred with native speakers of English

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

29. How much of your use of English during your school years occurred with non-native speakers of English

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

30. To what extent were your English classes at school conducted

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) by a native English speaker or by a teacher who has English as one of his/her native languages? (e.g. Greek-English bilingual)						
b) by a Greek(-Cypriot) native speaker?						
c) by a native speaker of another language?						
d) in English?						
e) in Greek?						

31. How often do you use (speak or listen to) English during your university years

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

32. How much of your use of English during your university years occurs with native speakers of English

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

33. How much of your use of English during your university years occurs with non-native speakers of English

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) in class?						
b) at home?						
c) in social settings?						

34. In your opinion, how important is it to learn English?

(Not important at all) 1 2 3 4 5 6 (Very important)

35. How important do you consider a good pronunciation of English to be?

(Not important at all) 1 2 3 4 5 6 (Very important)

36. Do you think that your teachers/tutors paid attention to the students' pronunciation in your English classroom (e.g. did they make corrections, did they teach specific sounds etc.)?

□ Yes □ No □ Maybe □ In some cases (e.g. at language institutes)

37. What did they do? (Select all that apply)

□ Teaching the pronunciation of specific sounds
□ Teaching the pronunciation of words
□ Teaching the pronunciation of phrases and sentences
□ Correcting students' pronunciation
□ Listening to recordings asking students to pay attention to pronunciation
□ Listening to recordings regularly (at least once a week)
□ Watching videos or movies regularly (at least once a week)
□ None of the above
□ Other_____
_____

38. In the past, I have practised my pronunciation skills…

	Never	Rarely	Sometimes	Often	Almost all the time	All the time
a) on my own						
b) while watching videos						
c) while watching movies or TV						
d) while listening to recordings for pronunciation improvement						
e) by observing the speech of English-speaking relatives or friends						
f) by asking someone to help me practise						
g) through pronunciation training						
h) in any other way						

39. When communicating in English, it is more important (Select all that apply):

□ for people to understand what I want to say
□ to have a good or native-like accent
□ I do not think my accent is important
□ other _____

40. Do you think having a good English pronunciation has any positive effects in any aspect of your professional or personal life? □ Yes □ No □ Maybe

41. If yes, please state in which aspects you believe it may be helpful to have a good English pronunciation. If no, please state why it is irrelevant to you.

_____

_____

_____

42. How much do you try to improve your English pronunciation?

(Not at all) 1 2 3 4 5 6 (A lot)

43. How confident are you in speaking English?

(Not at all) 1 2 3 4 5 6 (A lot)

44. My pronunciation level in English… (Select all that apply)

□ makes me uncomfortable when speaking English
□ makes me feel worried that I might be misunderstood
□ makes me feel worried about career opportunities
□ can meet the expectations of employers
□ negatively affects my confidence in speaking English
□ positively affects my confidence in speaking English
□ does not need to be improved
□ definitely needs improvement
□ Other: _____

_____

45. I participate in this study… (Select all that apply)

□ to try and improve my pronunciation
□ to see how pronunciation training is carried out
□ to see if it is possible to change or improve a person's pronunciation
□ because it will improve my pronunciation without much effort
□ but I don't believe that I can improve my pronunciation
□ out of curiosity
□ Other: _____

_____

46. Do you consider yourself as having normal hearing and no speech impairments?
□ Yes □ No □ Prefer not to say

47. Do you have any other comments?

_____
_____
_____

48. Would you like to be considered for a follow up interview? □ Yes □ No

49. Would you like to receive a summary of the results upon completion of the study?

□ Yes □ No

50. Please give us your email address if you would like to be considered for a follow up interview, and/or if you would like to receive a summary of the results upon completion of the study. Alternatively, if you prefer to be contacted via telephone, please provide your phone number: _____

Thank you for taking the time to complete this questionnaire.

If you have any questions, please feel free to contact the researcher:

Dimitra Dimitriou
University of Central Lancashire
E-mail address: dimitriou.dimitra18@gmail.com

B.2 Questionnaire completed by NE participants

Linguistic Background Questionnaire

Please fill in this questionnaire in full and as honestly as possible. Please note that you may ask the researcher for clarifications if you feel that a question is confusing.

It should not take more than 10 minutes to complete the questionnaire. Thank you in advance for taking the time to complete this questionnaire.

1. Name (Code): _____

2. Gender: □ Male □ Female □ Prefer not to say

3. Age: _____

4. Year of Studies: □ 1st □ 2nd □ 3rd □ 4th □ PG student □ Other

5. Degree Title: _____

6. Do you believe that you have a British Southern accent? □ Yes □ No □ Maybe □ Not sure

7. Were you born and brought up in a Southern region of England? □ Yes □ No

8. Do you consider yourself as a bilingual or a multilingual? □ Yes □ No

9. If yes, please state your native languages: _____

10. If you left the South and lived somewhere else for a period longer than 1 year, please state below the place where you stayed and for how long:

_____
_____
_____

11. English is my native language as well as that of my parents: □ Yes □ No

12. I have always used English as my main language at home and in social settings:

□ Yes □ No

13. If your answer was no, please explain (what other languages did you use, when, and why):_____
_____
_____

14. Were your parents born and brought up in a Southern region of England? □ Yes □ No

15. If your answer was no, please state the place where your parents were born and/or brought up, as well as their age of arrival in a Southern region of England.

_____
_____
_____

16. Please state any other languages you speak apart from English, the age you started learning them, and your proficiency in each language (Beginner, Intermediate, Advanced, Near-native):

_____
_____
_____

17. Approximately how often do you have conversations with speakers whose native language is not English and have foreign-accented speech? (1-Never, 6-All the time)

□ 1 □ 2 □ 3 □ 4 □ 5 □ 6

18. If you encounter second-language learners with a foreign accent in your daily life, which aspects of their speech make it difficult for you to understand them?

□ Errors is grammar and syntax
□ Disfluency (e.g. regular pauses, fillers such as "erm, um…")
□ Intonation (tone, pitch-range, loudness, rhythmicality, tempo)
□ Pronunciation of certain vowels
□ Pronunciation of certain consonants
□ General pronunciation (e.g. of words)
□ False starts and corrections
□ None of the above
□ Other: _____
_____

19. Do you have any other comments?

_____
_____
_____

20. Would you like to receive a summary of the results upon completion of the study?
□ Yes □ No

21. If yes, please give us your email address: _____
_____

Thank you for taking the time to complete this questionnaire.

If you have any questions, please feel free to contact the researcher:

Dimitra Dimitriou
University of Central Lancashire
E-mail address: dimitriou.dimitra18@gmail.com

B.3 Questionnaire completed by CYG trainees after training

Post-test Questionnaire – Evaluation

Please fill in this questionnaire in full and as honestly as possible. Please note that you may ask the researcher for clarifications if you feel that a question is confusing.

It should not take more than 15 minutes to complete the questionnaire. Thank you in advance for taking the time to complete this questionnaire.

1. Name (Code): _____

2. Do you think that the training helped you improve your pronunciation?

□ Yes □ No □ I don't know

3. If yes, do you think that it helped you improve your pronunciation of (Choose all that apply):

□ Vowels □ Words □ Sentences □ General pronunciation

4. Do you believe that this type of training can be beneficial for you in any way in the future?

□ Yes □ No □ I don't know

5. If yes, in what ways? (Choose all that apply)

□ Improving your career prospects

□ Making you more understandable when you speak English

□ Increasing your chances of visiting or moving to an English-speaking country

□ Increasing your chances of studying at an English-speaking country or an English-speaking university

□ Improving your academic performance

□ Increasing your confidence in speaking English

□ Improving your personal and/or professional life in various ways

□ Making you sound more professional

□ Other (Please state):_____

_____

_____

6. Do you think that you are more motivated to try and improve your pronunciation after this training?

□ Yes □ No □ Maybe

7. Are you more aware of your pronunciation difficulties after this training?

□ Yes □ No □ I don't know □ Maybe

8. Do you think that you will pay more attention to your pronunciation and the pronunciation of others after this training?

□ Yes □ No □ Maybe

9. Please let us know your thoughts about the training (Choose all that apply):

□ Interesting

□ Boring

□ Too long in total

□ More sessions would have been more beneficial

□ The sessions were too long

□ Longer sessions would have been more beneficial

□ Too many vowels during one session

□ It was helpful that all vowels were included in each session

□ Generally helpful

□ Could have had more sessions in one week

□ It was tiring

□ Other (please state):_____

_____

10. Do you have any suggestions for the improvement of the training in terms of its length? (consider length of training as a whole or length of each session)

_____

_____

_____

11. Do you have any suggestions for the improvement of the training in terms of its intensity? (e.g. too many items during one session, or number of sessions per week)

_____

_____

_____

12. Do you have any suggestions for the improvement of the training in terms of the training items? (the sessions focused on vowels only)

_____

_____

_____

13. Do you have any suggestions for the improvement of the training in terms of the training stimuli? (i.e. the words used, the use of both real and nonsense words)

_____

_____

_____

14. Do you have any suggestions for the improvement of the training in terms of the accent used in the training? (a southern accent of England was used)

_____

_____

_____

15. Which was the easiest and most difficult group of sounds for you? Please rank them (1-easiest, 5-most difficult):

bit – beat bart – bat – but boot – butch bet – burt bot – bought

[ ] [ ] [ ] [ ] [ ]

16. Do you believe that this type of training can be beneficial for learners of English?

□ Yes □ No □ I don't know □ Yes, if the improvements suggested are considered

17. If yes, in what ways can it be beneficial for learners of English? (Choose all that apply)

□ Improving their career prospects

□ Making them more understandable when they speak English

□ Increasing their chances of visiting or moving to an English-speaking country

□ Increasing their chances of studying at an English-speaking country or an English-speaking university

□ Improving their academic performance

□ Increasing their confidence in speaking English

□ Improving their personal and/or professional lives in various ways

□ Making them sound more professional

□ Other (Please state):

_____

_____

Thank you for taking the time to complete this questionnaire.

If you have any questions, please feel free to contact the researcher:

Dimitra Dimitriou
University of Central Lancashire
E-mail address: dimitriou.dimitra18@gmail.com

B.4 Semi-structured interview questions

1. Can you tell me a bit about your relationship with relatives who live in English-speaking countries? (e.g. how often do they visit you, how often do you visit them, how much time do you spend together, do you speak English or Greek when you are together)
2. How often do you travel to English-speaking countries and for how long?
3. Talk to be about your school years:
    a. Do you remember any occasions where you spoke English with friends/classmates/acquaintances, at home, in social life and at school?
    b. Can you give a brief description of your experience in learning English at school? Consider your teachers, the language of instruction, the syllabus etc.
    c. To what extent did you have other opportunities to use or listen to English? E.g. watching movies, youtube videos etc.
    d. During these experiences using English, can you recall to what extent you used English with native and non-native speakers?
4. Moving on to your university years:
    a. Do you still have any English lessons at university? How are they, compared to your school years? Think about your teachers, the language of instruction, the syllabus etc.
    b. Do you have any comments about your experience at the university as concerns the use of English in the classroom? Do you find it difficult, do you prefer to participate in the classroom using Greek, do you avoid participating because you have to speak English? (remind anonymity)
    c. To what extent do you speak English at the university with friends/ classmates/acquaintances, at home or in social life?
    d. To what extent did you have other opportunities to use or listen to English? E.g. watching movies, YouTube videos etc.
    e. During these experiences using English, can you recall to what extent you used English with native and non-native speakers?
5. Did you have any teachers/tutors at any point during your education that paid attention to the students' pronunciation? If yes, can you give me some examples of what they did (e.g. did they make corrections, did they teach specific sounds, did they include any listening tasks asking you to pay attention to pronunciation, or did they normally ask you to watch videos or movies)?
6. What is your opinion about having a good pronunciation in English? Is there any way that your English pronunciation affects your current or future personal or professional life?
7. Have you ever tried to practise your pronunciation skills? If yes, in what ways?
8. Some people believe that it is more important for people to understand what they say even if they have a foreign accent, and others believe that sounding native-like is equally important. What is your opinion about this disagreement?
9. Are you confident in speaking English? If you are not, is this because of your pronunciation?

10. Do you ever feel uncomfortable speaking English with native English speakers, other Greek-Cypriots or any other non-native speakers of English?
11. Do you ever feel worried about your pronunciation? For example, are you worried that people might not be able to understand you, or that it can cost you a good career opportunity?
12. Do you think your pronunciation level is satisfactory and can therefore meet the expectations of employers?
13. Is there any particular future goal that you have which may require a good level of English? E.g. studying abroad, getting a job requiring high communication skills in English etc.
14. Why did you want to participate in this study? What are your expectations about it? Do you think it can help you improve your pronunciation? Do you believe this is possible?
15. If you were to see some improvements of your pronunciation after some training, would you feel more determined to try and improve your pronunciation even more?
16. Would you like to improve your pronunciation, or is this something that doesn't really concern you?
17. Do you think that the training helped you improve your pronunciation in any way? (e.g. in vowels, words, sentences, general pronunciation)
18. Do you think this type of training can be beneficial for learners of English? If yes, in what ways? Could we incorporate it in schools, afternoon institutes, for adult programmes, for individual learning?
19. Have you got any other comments?

Appendix C

C.1 Praat Script used for creating synthetic stimuli

```
# Matthew Winn
# August 2014
form Input Enter specifications for Formant settings
   comment shortest duration (ms):
   real shortdurms 60
   comment longest duration (ms):
   real longdurms 300
   comment how many steps in the continuum?
   integer steps 5
   comment enter minimum pitch
   real minpitch 70
   comment enter maximum pitch
   real maxpitch 300
   comment enter duration name prefix
   word durPrefix _dur_
endform
clearinfo
shortdur = shortdurms/1000
longdur = longdurms/1000
call printHeader
call makeContinuum steps shortdur longdur dur_ 1
call enumerateSounds
call identifyLandmarks
for thisSound from 1 to numberOfSelectedSounds
   select sound'thisSound'
        name$ = selected$("Sound")
        call makeDurationContinuum 'name$' start end shortdur longdur steps 'durPrefix$'
endfor
procedure makeDurationContinuum .name$ .start .end .shortdur .longdur .steps .suffix$
        select Sound '.name$'
        .endTime = Get end time
        To Manipulation... 0.01 minpitch maxpitch
        Extract duration tier
        for thisStep from 1 to .steps
                ratio = (dur_'thisStep')/(.end - .start)
                select DurationTier '.name$'
                        Remove points between... 0 .endTime
                        Add point... (.start-0.0001) 1
                        Add point... .start ratio
                        Add point... .end ratio
                        Add point... (.end+0.0001) 1
                select Manipulation '.name$'
                plus DurationTier '.name$'
                Replace duration tier
```

```
                select Manipulation '.name$'
                Get resynthesis (overlap-add)
                Rename... '.name$''.suffix$''thisStep'
        endfor
        select Manipulation '.name$'
        plus DurationTier '.name$'
        Remove
endproc
procedure identifyLandmarks
        select sound1
        firstName$ = selected$("Sound")
        Edit
                editor Sound 'firstName$'
                # prompts user to click on vowel beginning and end, create variables with
values at points clicked
                 pause Click Get start of segment to be manipulated, click Continue when
done
                 Move cursor to nearest zero crossing
                 start = Get cursor
                 pause Click Get end of segment to be manipulated, click Continue when done
                 Move cursor to nearest zero crossing
                 end = Get cursor
        Close
        endeditor
endproc
procedure enumerateSounds
        pause select all sounds to be used for this operation
        numberOfSelectedSounds = numberOfSelected ("Sound")
        for thisSelectedSound to numberOfSelectedSounds
                sound'thisSelectedSound' = selected("Sound",thisSelectedSound)
        endfor
endproc
procedure printHeader
        # creates simple header
        print Step 'tab$' Duration 'tab$' 'newline$'
endproc
procedure makeContinuum .steps .low .high .prefix$ printvalues
        for thisStep from 1 to .steps
                temp = (('thisStep'-1)*('.high'-'.low')/('.steps'-1))+'.low'
                '.prefix$''thisStep' = temp
                check = '.prefix$''thisStep'
                if printvalues = 1
                print '.prefix$''thisStep''tab$''check:2' 'newline$'
                endif
        endfor
endproc
```

C.2 Praat Script used for extracting vowel duration and formant values

This script opens each sound file in a directory, looks for a corresponding TextGrid in (possibly) a different directory, and extracts f0, F1, and F2 from the midpoint(s) of any labelled interval(s) in the specified TextGrid tier. It also extracts the duration of the labelled interval(s). All these results are written to a tab-delimited text file.
The script is a modified version of the script "collect_formant_data_from_files.praat" by Mietta Lennes, available here: http://www.helsinki.fi/~lennes/praat-scripts/
The modifications were done by Dan McCloy (drmccloy@uw.edu) in December 2011.

This script is distributed under the GNU General Public License.
Copyright 4.7.2003 Mietta Lennes

```
form Get pitch formants and duration from labeled segments in files
	comment Directory of sound files. Be sure to include the final "/"
	text sound_directory /home/dan/Desktop/sound files/
	sentence Sound_file_extension .wav
	comment Directory of TextGrid files. Be sure to include the final "/"
	text textGrid_directory /home/dan/Desktop/text grids/
	sentence TextGrid_file_extension .TextGrid
	comment Full path of the resulting text file:
	text resultsfile /home/dan/Desktop/resultsfile.txt
	comment Which tier do you want to analyze?
	integer Tier 1
	comment Formant analysis parameters
	positive Time_step 0.01
	integer Maximum_number_of_formants 5
	positive Maximum_formant_(Hz) 5500
	positive Window_length_(s) 0.025
	real Preemphasis_from_(Hz) 50
	comment Pitch analysis parameters
	positive Pitch_time_step 0.01
	positive Minimum_pitch_(Hz) 75
	positive Maximum_pitch_(Hz) 300
endform

# Make a listing of all the sound files in a directory:
Create Strings as file list... list 'sound_directory$'*'sound_file_extension$'
numberOfFiles = Get number of strings

# Check if the result file exists:
if fileReadable (resultsfile$)
	pause The file 'resultsfile$' already exists! Do you want to overwrite it?
	filedelete 'resultsfile$'
endif

# Create a header row for the result file: (remember to edit this if you add or change the analyses!)
```

```
header$ = "Filename  TextGridLabel  duration        f0_midpoint    F1_midpoint
        F2_midpoint'newline$'"
fileappend "'resultsfile$'" 'header$'

# Open each sound file in the directory:
for ifile to numberOfFiles
        filename$ = Get string... ifile
        Read from file... 'sound_directory$''filename$'

        # get the name of the sound object:
        soundname$ = selected$ ("Sound", 1)

        # Look for a TextGrid by the same name:
        gridfile$ = "'textGrid_directory$''soundname$''textGrid_file_extension$'"

        # if a TextGrid exists, open it and do the analysis:
        if fileReadable (gridfile$)
                Read from file... 'gridfile$'

                select Sound 'soundname$'
                To Formant (burg)... time_step maximum_number_of_formants
maximum_formant window_length preemphasis_from

                select Sound 'soundname$'
                To Pitch... pitch_time_step minimum_pitch maximum_pitch

                select TextGrid 'soundname$'
                numberOfIntervals = Get number of intervals... tier

                # Pass through all intervals in the designated tier, and if they have a label, do
the analysis:
                for interval to numberOfIntervals
                        label$ = Get label of interval... tier interval
                        if label$ <> ""
                                # duration:
                                start = Get starting point... tier interval
                                end = Get end point... tier interval
                                duration = end-start
                                midpoint = (start + end) / 2

                                # formants:
                                select Formant 'soundname$'
                                f1_50 = Get value at time... 1 midpoint Hertz Linear
                                f2_50 = Get value at time... 2 midpoint Hertz Linear

                                # pitch:
                                select Pitch 'soundname$'
```

```
                        f0_50 = Get value at time... midpoint Hertz Linear

                        # Save result to text file:
                        resultline$ = "'soundname$'   'label$''duration'        'f0_50'
    'f1_50' 'f2_50''newline$'"
                        fileappend "'resultsfile$'" 'resultline$'

                        # select the TextGrid so we can iterate to the next interval:
                        select TextGrid 'soundname$'
                    endif
                endfor
                # Remove the TextGrid, Formant, and Pitch objects
                select TextGrid 'soundname$'
                plus Formant 'soundname$'
                plus Pitch 'soundname$'
                Remove
        endif
        # Remove the Sound object
        select Sound 'soundname$'
        Remove
        # and go on with the next sound file!
        select Strings list
endfor

# When everything is done, remove the list of sound file paths:
Remove
```

Appendix D

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**21.3**	**61.3**	15							2.5	
FLEECE	**13.3**	**82.7**	2.7							1.3	
DRESS	2.5		**87.5**	5		1.3	1.3	1.3		1.3	
NURSE			**16.9**	**72.3**	1.5					6.2	3.1
BATH				8	**24**	16	10	**40**	2		
TRAP			3.8	6.3	5	**68.8**	3.8	**12.5**			
STRUT			5.3	9.3	9.3	**41.3**	9.3	24	1.3		
LOT				1.3			6.7	**72**	**14.7**	2.7	2.7
NORTH				1.3			4	**56**	**32**	2.7	4
FOOT				2.5				1.3	2.5	**52.5**	**41.3**
GOOSE								1.7	5	**31.7**	**61.7**

Table D1. Confusion matrix of NE raters' percentage of responses to CYG trainees' productions of each target vowel at T1 in the elicitation task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**21.7**	**76.7**									1.7
FLEECE	10	**83.3**		1.7							5
DRESS		1.7	**90**	3.3				1.7	1.7		1.7
NURSE	10		10	**70**			5	5			
BATH				2.5	35	**12.5**	5	**37.5**	7.5		
TRAP					**16.4**	**60**	7.3	**16.4**			
STRUT					14	**52**	8	**24**	2		
LOT					1.8		3.6	**56.4**	**34.5**	1.8	1.8
NORTH							3.3	**58.3**	**31.7**	1.7	5
FOOT							1.7	1.7		**50**	**46.7**
GOOSE				2.2			2.2	2.2	2.2	**37.8**	**53.3**

Table D2. Confusion matrix of NE raters' percentage of responses to CYG controls' productions of each target vowel at T1 in the elicitation task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**16.3**	**82.5**					1.3				
FLEECE	**14.7**	**80**	4		1.3						
DRESS			**90**	6.3		1.3				2.5	
NURSE			6.7	**90**				3.3			
BATH				5.7	**52.9**	11.4	1.4	**25.7**	2.9		
TRAP					11.3	**58.8**	1.3	**26.3**	2.5		
STRUT			1.4		5.4	**60.8**	1.4	**28.4**	2.7		
LOT				2.7		1.3	1.3	**65.3**	**26.7**	1.3	1.3
NORTH	1.3			1.3		1.3	2.7	**38.7**	**50.7**	2.7	1.3
FOOT										**60**	**40**
GOOSE	1.4						1.4		5.7	**31.4**	**60**

Table D3. Confusion matrix of NE raters' percentage of responses to CYG trainees' productions of each target vowel at T1 in the wordlist-reading task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	8.3	**88.3**									3.3
FLEECE	7.3	**87.3**			1.8						3.6
DRESS		1.7	**86.7**	10			1.7				
NURSE			**60**	**35**	2.5	2.5					
BATH					**30**	**41.7**	5	23.3			
TRAP				1.8	16.4	**58.2**	5.5	**18.2**			
STRUT				1.8	18.2	**47.3**	10.9	**21.8**			
LOT					1.7	1.7		**61.7**	**35**		
NORTH					1.8		1.8	**45.5**	**47.3**	1.8	1.8
FOOT				5						**60**	**35**
GOOSE									3.6	**40**	**56.4**

Table D4. Confusion matrix of NE raters' percentage of responses to CYG controls' productions of each target vowel at T1 in the wordlist-reading task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**26.7**	**53.3**	17.3	1.3							1.3
FLEECE	**26.3**	**66.3**	3.8		1.3		1.3				1.3
DRESS	10		**76.3**	**11.3**			2.5				
NURSE	5.7	1.4	**12.9**	**77.1**			1.4		1.4		
BATH				10.8	**44.6**	12.3		**23.1**	9.2		
TRAP			1.3	5.1	6.3	**63.3**	**17.7**	6.3			
STRUT			2.7	2.7	8	**53.3**	13.3	**17.3**	2.7		
LOT				2.7	1.4	2.7	6.8	**55.4**	**24.3**	2.7	4.1
NORTH			1.3	2.7	2.7		6.7	**41.3**	**34.7**	5.3	5.3
FOOT	1.3		1.3				10			**46.3**	**41.3**
GOOSE	1.3	1.3					2.7	5.3	9.3	**21.3**	**58.7**

Table D5. Confusion matrix of NE raters' percentage of responses to CYG trainees' productions of each target vowel at T2 in the elicitation task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**16.7**	**81.7**	1.7								
FLEECE	**27.3**	**72.7**									
DRESS	3.6		**85.5**	**10.9**							
NURSE			**36**	**60**		4					
BATH				2.5	**27.5**	**27.5**	5	**32.5**	2.5	2.5	
TRAP				3.3	10	**50**	10	**26.7**			
STRUT				2.2	**28.9**	**40**	2.2	26.7			
LOT				1.7	1.7		1.7	**65**	**28.3**	1.7	
NORTH					4.4			**62.2**	**22.2**	8.9	2.2
FOOT							1.7	3.3		**48.3**	**46.7**
GOOSE				2.2			2.2			**53.3**	**42.2**

Table D6. Confusion matrix of NE raters' percentage of responses to CYG controls' productions of each target vowel at T2 in the elicitation task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**12**	**86.7**									1.3
FLEECE	10	**88.8**								1.3	
DRESS	2.5		**85**	**12.5**							
NURSE			1.7	**95**		1.7	1.7				
BATH				6.3	**62.5**	6.3		**20**	3.8	1.3	
TRAP					10	**63.8**	3.8	**20**	2.5		
STRUT				2.5	7.5	**57.5**	3.8	**25**	2.5	1.3	
LOT				1.3	1.3		5.3	**64**	**21.3**	6.7	
NORTH				1.3			1.3	**41.3**	**54.7**	1.3	
FOOT				3.1			1.5	3.1		**56.9**	**35.4**
GOOSE					1.3			1.3	9.3	**25.3**	**62.7**

Table D7. Confusion matrix of NE raters' percentage of responses to CYG trainees' productions of each target vowel at T2 in the wordlist-reading task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

RESPONSE / TARGET VOWEL	KIT	FLEECE	DRESS	NURSE	BATH	TRAP	STRUT	LOT	NORTH	FOOT	GOOSE
KIT	**14.5**	**83.6**	1.8								
FLEECE	**10.9**	**85.5**	3.6								
DRESS		1.7	**86.7**	8.3	3.3						
NURSE			**71.1**	**28.9**							
BATH				1.7	**30**	**40**	1.7	26.7			
TRAP					**29.1**	**56.4**	1.8	12.7			
STRUT		1.8	1.8		**20**	**54.5**	5.5	16.4			
LOT					1.8		3.6	**60**	**25.5**	7.3	1.8
NORTH								**52**	**36**	8	4
FOOT									2.9	**82.9**	**14.3**
GOOSE								3.6	7.3	**45.5**	**43.6**

Table D8. Confusion matrix of NE raters' percentage of responses to CYG controls' productions of each target vowel at T2 in the wordlist-reading task. Blank: responses below 1%, underlined: predicted confusion patterns, bold: two most common responses for each vowel (when above 10%)

Appendix E

E.1 Interview with CYG01 (translated from CYG)

DD: So, let me start the recording. OK, so you are participant CYG01. Are you aware that this interview is being recorded? Do you agree to that?

CYG01: Yes.

DD: Good. So… According to your questionnaire, I can see here that you have relatives who live in England, in the UK. I just wanted to ask you if you can tell me more about your relationship with them, that is, how often do you see them, how often do they come or you go, how much time do you spend together, if you speak Greek or English when you are together…?

CYG01: ...

DD: Do they come to Cyprus at all?

CYG01: Yes. They do come to Cyprus.

DD: How often approximately?

CYG01: Once-twice per year.

DD: OK. Approximately for how long do they come, a week, a month, how…

CYG01: Approximately one week, two weeks.

DD: Do you spend time with them?

CYG01: Very little.

DD: Very little. Approximately how much that is? One whole day, one hour…

CYG01: One day.

DD: One day. OK. Approximately one day, that is twice a year? One day every time they come?

CYG01: Yes, yes.

DD: Do you go there at all?

CYG01: No.

DD: Not at all. OK. And when you are together, do you speak Greek or English?

CYG01: Greek.

DD: Greek. Good. And you said you don't visit them. Have you visited any other English-speaking countries in the past?

CYG01: No.

DD: No, not ever?

CYG01: No.

DD: OK, so, I'll take you back to your school years. So, when you were at school, that is middle school and high school, do you remember any times when you used to speak English, whether it was with friends, with classmates, with acquaintances, with people you didn't know… Did you speak English at all during your school years?

CYG01: Only in private lessons.

DD: Only in private lessons. OK, how often did you have private lessons?

CYG01: Twice a week. Only.

DD: Only that. For 1 hour each time? 2 hours a week?

CYG01: Yes.

DD: OK. Your experience when you were learning English at school or in afternoon lessons, how was it? Can you… Think about your teachers, how did they speak, in what language, what were you doing in the lesson for example? What was in the syllabus… as you remember it. The image you have in your mind, your experience.

CYG01: I don't remember.

DD: You don't remember? Did you use to speak in Greek or English in class?

CYG01: In my English lessons?

DD: Yes.

CYG01: Sometimes in Greek, sometimes in English.

DD: OK, is that about 50-50 let's say?

CYG01: Yes.

DD: OK. And your syllabus, did you do grammar, vocabulary…?

CYG01: Everything.

DD: Everything. Did you do listening activities too?

CYG01: Yes.

DD: Speaking as well?

CYG01: Yes.

DD: And your teachers, did they use to make any comments on your… on your pronunciation, that is, telling you how to pronounce something, correcting you…

CYG01: Not to me personally.

DD: Not to you personally, but in general, did they do something like that to anyone?

CYG01: Yes.

DD: OK, did you have, let's say some time during the lesson when, for example, they told you, for these 10 minutes we will be talking about the pronunciation of these words? Did you have some allocated time in class when you discussed stuff like that?

CYG01: During speaking activities yes.

DD: During speaking activities. OK. OK. During your school years, did you have any other opportunities to speak or listen to English? That is, did you use to watch movies, Youtube videos, podcasts…? During your school years, did you use to play computer games?

CYG01: Yes. Yes.

DD: What did you use to do?

CYG01: I used to watch series… As you said now. I used to watch series, movies, Youtube, Netflix…

DD: OK, gaming?

CYG01: Yes.

DD: OK. Let's say about the movies, the series, how many hours, for example, did you spend watching movies?

CYG01: I spent half of my day on computer games, movies…

DD: OK, half of your day, that is your afternoon I imagine…

CYG01: Yes, yes, in the afternoon.

DD: How many hours, approximately?

CYG01: ...

DD: You finished school, did your homework… Shall we say from 5 until 10?

CYG01: Yes, yes…

DD: So you could say you used to do that for 4 hours a day? 5 hours a day?

CYG01: 4-5, yes.

DD: 4-5 hours. OK. That is half of your day, half of your afternoon.

CYG01: Yes.

DD: OK. And you mentioned movies, series, videos…

CYG01: Gaming…

DD: And gaming, but in computer games, you only write things to each other, you do not speak to each other, right?

CYG01: Yes.

DD: Just writing. OK. During all of these that you mentioned, that is at school, in class when you spoke in English and in the movies you watched etc., approximately how much of that was with native speakers and how much was with non-native speakers?

CYG01: Wait, in movies?

DD: In your whole experience. Divide it into native and non-native speakers. How much contact did you have with native speakers, and how much with non-native speakers?

CYG01: Movies had native speakers. In computer games, with non-native sp… From all…

DD: Other non-native speakers… yes… In class, did you have any teachers who were native speakers?

CYG01: Most of them, yes… most of them were…

DD: English?

CYG01: No, Cypriots.

DD: So most of them were non-native speakers.

CYG01: Yes.

DD: OK. Most of them were Cypriots. Did you have any teachers who were native speakers? English?

CYG01: No.

DD: No?

CYG01: No.

DD: OK. And the series you used to watch were with native speakers but I imagine from different regions. Were they mostly American speakers, were they mostly British speakers, Australian, if you can tell them apart?

CYG01: We can tell them apart but... it depends, I used to watch a bit of everything… in terms of movies…

DD: A variety.

CYG01: A variety, yes.

DD: OK, you didn't mind if it was with…

CYG01: Yes, yes.

DD: …any accent. OK. Moving on to the university. During your university years, more recently. So, do you have any English lessons at the university? Not lessons in English, lessons for English, to learn English.

CYG01: During the first year and… I think we only had English lessons during the first year.

DD: OK, and how were these lessons compared to the private lessons you had during your school years? Did they speak in English, were your teachers English, were they Cypriots, the language you used, did you speak Greek at all, was your syllabus approximately the same as with school or did it have any additional things?

CYG01: ...

DD: Let's start with the teachers.

CYG01: Yes.

DD: What…

CYG01: Only English with teachers.

DD: You only spoke in English.

CYG01: Yes.

DD: OK.

CYG01: Because of the other students who are…

DD: OK. Were your teachers English or Cypriot, as far as you know?

CYG01: Cypriots.

DD: Cypriots. Was there anyone who was a foreigner? Speaking another language?

CYG01: ...

DD: Or were they Cypriots?

CYG01: Cypriots.

DD: Cypriots. OK. You only spoke in English… And your syllabus, was it let's say similar to that of school, grammar, listening, speaking…?

CYG01: ...

DD: As far as you can remember.

CYG01: We had listening activities… and…

DD: In general, did they pay any attention to your speaking, having you speak…?

CYG01: No, because we were admitted into an English university, I think it's… it's reasonable to assume that we know how to speak English…

DD: OK…

CYG01: ...Since…

DD: It goes without saying that you speak it.

CYG01: Yes, yes.

DD: So you did not speak English in class?

CYG01: We did, due to the fact that it's an English university.

DD: I mean you didn't have any speaking activities, that is, to have a topic and to have to expand on it… or do a presentation or something?

CYG01: ...

DD: Do you not remember?

CYG01: No.

DD: You don't remember. OK. So it's possible that you didn't speak at all in class?

CYG01: No I did speak…

DD: You did speak.

CYG01: English.

DD: English. You were answering the teacher's questions.

CYG01: Yes, yes.

DD: OK. At any stage during your education, at school, at the university, did you have any teachers who paid attention to your pronunciation, that is to tell you "this is not pronounced like this, but like this", or who demonstrated… or who gave you examples, or corrected you, or taught you specific sounds as with our training, or who gave you specific sounds to pronounce…?

CYG01: In private lessons.

DD: In private lessons.

CYG01: Yes.

DD: What did they do? Did they give you listening activities, asking you to "listen to the pronunciation of this specifically", did they ask you to watch many videos all the time so that

you would get the pronunciation… what did they do? What did they do to maybe help you with pronunciation?

CYG01: We used to watch a lot of Ted Talks…

DD: OK…

CYG01: ...

DD: So, did they use to ask you to pay attention to pronunciation?

CYG01: Yes.

DD: OK.

CYG01: Yes… When we said something wrong they used to correct us…

DD: How did they correct you, with their own examples, or did they play the words online?

CYG01: They repeated it correctly.

DD: They repeated it correctly. OK.

CYG01: Yes.

DD: OK. And did you use to have anything like our training, e.g. taking a group of sounds and telling you that these are pronounced like this and that?

CYG01: Not necessarily…

DD: Or any activities for pronunciation specifically?

CYG01: No, as far as I can remember.

DD: OK. So, your opinion on pron… On having a good pronunciation in English. What is it? Do you think it's important, do you think that it doesn't matter, do you think that… are you neutral…? What do you think about having a good pronunciation in English?

CYG01: Pronunciation… Basically, because there are various accents… I thi… I don't know if there is a correct pronunciation or not. Because there are… many accents, I don't know how to say it.

DD: Yes, correct, yes.

CYG01: OK.

DD: Do you think it is important to speak with one of those accents, any one of them? That is, do you try to imitate one of them because it's good to…

CYG01: Yes, definitely some of them are wrong accents, let's say for example, us Cypriots…

DD: OK, so that is, us Cypriots, we speak wrongly.

CYG01: Yes.

DD: OK. But I mean the accent of Americans, of English people, from England, of Australians… Do you think that it's good for someone who learns to speak English to speak it with one of those accents?

CYG01: Yes.

DD: Is this important?

CYG01: Yes.

DD: OK. Why? Why would someone want to sound like English is their first language let's say?

CYG01: Not to sound like it's their first language, but to… say that they speak correctly.

DD: To speak correctly.

CYG01: Yes, to not speak it wrongly let's say…

DD: OK. Not to speak wrongly. Yes, fine. OK. Do you think that people won't be able to understand you if you don't speak with one of those accents?

CYG01: With one of those accents maybe… if you have one of those accents they may not be able to understand you anyway, so…

DD: OK. So do you think it's a good thing to speak with a good accent, a correct English accent?

CYG01: Yes.

DD: Is there any one of them you prefer? American, British…?

CYG01: The Scottish accent!

DD: Scottish!

CYG01: I like it, yes!

DD: OK. But you think that if you have it, other English people will not be able to understand you.

CYG01: Certainly.

DD: OK. Do you think that having a good accent affects your personal or professional life in any way?

CYG01: No.

DD: No. So whether you speak with a Cypriot accent or an English accent does not affect… I'm not saying it's wrong, I'm just putting it in a different way so that you can tell me…

CYG01: I get what you are trying to say but…

DD: There is no right or wrong answer, it's your opinion.

CYG01: …it doesn't play a role, how to put it… I think pronunciation doesn't have a role to play on how… others see you, to put it this way.

DD: OK. Where does pronunciation have a role to play?

CYG01: In… self-improvement.

DD: For yourself, that is.

CYG01: Yes.

DD: …to do it for yourself.

CYG01: Yes.

DD: OK. In professional life, let's say, what do you think has a role to play, in relation to English?

CYG01: If you want to get a job, knowing how to speak the language correctly will help a lot.

DD: OK. Right, but you mean your pronunciation doesn't matter.

CYG01: No, your pronunciation should be correct, to speak English correctly.

DD: Let me state it and you tell me if you agree. You mean for people to understand you.

CYG01: Yes, yes.

DD: To be comprehensible in what you say.

CYG01: Yes.

DD: So it's good to have a good pronunciation, but for your own use. Not because it will help you get a job easier…

CYG01: Look, when you speak correctly, with a correct pronunciation, it's even better for jobs as well.

DD: OK, when you say correct pronunciation, which pronunciation do you mean?

CYG01: Correct English, not Cypriot English let's say.

DD: OK. So the English that… of those people whose first language is English.

CYG01: Yes.

DD: Of English people.

CYG01: Yes.

DD: OK. So, apart from our training, have you ever practiced your pronunciation?

CYG01: No.

DD: No?

CYG01: No.

DD: Never?

CYG01: No.

DD: OK. And when you listen to English in movies, for example, you haven't tried, you've never said "how does he pronounce this, let me imitate it".

CYG01: No.

DD: Nothing, ever.

CYG01: No.

DD: OK. So, some people believe that it's more important… for others to understand what we want to say rather than… even if our pronunciation sounds foreign-accented. That is, it doesn't matter if we speak with the Cypriot accent, as long as the other person understands us. Some other people say that it's more… it's equally important to sound like English people, to sound native-like. What do you think? That it's equally important…

CYG01: I agree with the second.

DD: With the second? That it's equally important to have a good accent?

CYG01: Yes, yes.

DD: OK. So both speaking well so that others understand you, and sounding native-like as well.

CYG01: Yes.

DD: Irrespective of whether they understand you anyway, even with your own Cypriot accent let's say.

CYG01: Yes.

DD: Yes. OK. So you think pronunciation is very important.

CYG01: Yes.

DD: OK. So… But you've never practiced it.

CYG01: Yes.

DD: OK.

CYG01: Because as concerns me personally… when I see someone else, how to put it… they need to know how to speak correctly.

DD: Yes.

CYG01: With me, the subject doesn't matter.

DD: OK, for you it's too much work…

CYG01: Yes, exactly, too much effort.

DD: And you don't want to make the effort let's say.

CYG01: Exactly.

DD: You don't think it will help you that much at your own job.

CYG01: Exactly.

DD: OK. Good. This is very interesting. So, are you confident in speaking English?

CYG01: Yes.

DD: OK. So you are confident in speaking English.

CYG01: Yes…

DD: Very well.

CYG01: With foreigners.

DD: With foreigners! Interesting. Why? With which foreigners? With English people? With other foreigners?

CYG01: Yes.

DD: With both of these groups? As long as the other person is not Cypriot?

CYG01: Hmm… no, with foreigners…? I don't know… for example, at work when I used to speak with foreigners only, I was confident enough.

DD: OK. Yes.

CYG01: With Cypriots I don't know…

DD: You become self-conscious.

CYG01: Yes. This…

DD: OK.

CYG01: Not self-conscious… this is a strong word… self-conscious.

DD: Yes. What does… Do you only have this problem with Cypriots or when you speak with English people as well?

CYG01: ...

DD: Because we have 3 categories. We have foreigners, who do not speak, who are not English and they are foreigners, we have Cypriots, and we have English people who are foreigners.

CYG01: With… Look, with foreigners, I don't mind speaking. With English people, I don't mind speaking. With Cypriots, I don't speak.

DD: Do you not speak with Cypriots?

CYG01: Yes.

DD: In general, or do you not speak English?

CYG01: Yes.

DD: So do you mind speaking English with Cypriots?

CYG01: ...

DD: Do you feel weird, do you feel something, that…?

CYG01: I prefer to speak Greek with Cypriots.

DD: OK. If there's a third person with you and they are a foreigner? If there's 3 of you and 2 of you are Cypriots and the other is a foreigner…

CYG01: In that case… English.

DD: You don't mind in that case.

CYG01: Yes.

DD: OK.

CYG01: When we have to, when it's appropriate for all 3 of us to understand what…

DD: So basically, only when you are forced to speak.

CYG01: Yes. In that case I don't have a problem.

DD: If it's a person who's half-English half-Cypriot? And they understand both languages?

CYG01: In that case Greek.

DD: In that case Greek. So the language that you understand better, that is?

CYG01: Yes.

DD: OK. Good.

CYG01: Each person… in their language.

DD: OK, I've already asked the same question too. So basically, it's that you feel uncomfortable, let's say with Cypriots?

CYG01: No. Let me put it in a different way. It's that you feel more comfortable speaking…

DD: Speaking Greek.

CYG01: Yes, exactly.

DD: It's not that… you are embarrassed, let's say.

CYG01: No.

DD: It's not a matter of being embarrassed.

CYG01: No.

DD: So you wouldn't have a problem, it's just that you feel that it's easier, it takes less effort.

CYG01: Yes.

DD: OK. Good. Is there a time when you happen to feel… that you worry about your accent? For example, do you worry that someone might not be able to understand you or that you may want to get a good job and this particular issue with your pronunciation might hold you back in some way?

CYG01: In Cyprus… I think…

DD: I'm talking about your own accent.

CYG01: Yes.

DD: That is, do you feel that with your own accent, there is a job where they may not hire you due to your accent?

CYG01: No.

DD: No. Or do you think that someone might not be able to understand you due to your accent?

CYG01: Because we are talking about Cyprus now, no.

DD: Not in Cyprus. Yes. With foreigners? When you speak with foreigners, that they may not understand you?

CYG01: I would say about 80% of English people understand me.

DD: OK. So you've never felt that someone doesn't understand you.

CYG01: No, no.

DD: No. You are… you speak well enough for others to understand.

CYG01: Yes.

DD: So having a good pronunciation is something extra. It's not… You think that you are comprehensible…

CYG01: Yes, yes.

DD:...irrespective of whether you have the accent of an English person or not.

CYG01: Yes.

DD: OK. So… Then you also think, according to your answer, that your pronunciation level is satisfactory to meet the expectations of any employer.

CYG01: Exactly.

DD: In Cyprus, abroad, everywhere…?

CYG01: Yes.

DD: So everywhere, anywhere.

CYG01: Hmm... When we say everythere…

DD: I mean that even if you go to England to find a job, you think that your level is good enough.

CYG01: Yes, yes.

DD: Yes.

CYG01: Not to go to Africa, let's say.

DD: Yes, let's say to an English-speaking country, to go work in an English-speaking country, you think that the level is good enough…

CYG01: Good enough.

DD: OK. So do you have any immediate or future goal that would require you to have a good level of English, that is, do you want to go to England for a master's degree, or do you want to find a job that requires you to talk a lot like the one you previously had, where you need to speak English fluently? Do you have a goal which would require you to speak English fluently? Not that you don't, I mean…

CYG01: No.

DD: No. Are you thinking about studying in England?

CYG01: N…No.

DD: If you were to do, do you think it's satisfactory, your…

CYG01: Yes, yes.

DD: Yes. And… are you thinking… is there a job that you would like to do which would require… communication skills, and you would definitely need to speak English well? Is there any such job on your mind, which you would like to do?

CYG01: Yes; Yes.

DD: You tell me. Is there?

CYG01: Yes! Yes.

DD: And you would talk… let's say you would communicate a lot in English?

CYG01: If needed, yes.

DD: If needed. Not to be, let's say, an English teacher.

CYG01: Yes. I wouldn't be an English teacher.

DD: Or to be, let's say, a news broadcaster in England.

CYG01: No, no.

DD: It's not… You don't have any jobs in mind which would require… that much communication…

CYG01: No.

DD: OK. So it's not among your immediate goals to do anything that would require you to speak English perfectly let's say. OK.

CYG01: [nodding "No"]

DD: OK. Good. So, why did you want to participate in this study? What were you expecting when I told you about this study, and did you think at any point of the training that it could improve your pronunciation?

CYG01: ...

DD: Did you do it simply out of curiosity, did you think that it's an opportunity to see if your pronunciation can be improved, what did you think?

CYG01: Nothing, I just did it.

DD: You just did it. OK. Was it interesting for you? Or was it kind of a chore?

CYG01: Neither.

DD: You were neutral.

CYG01: I was neutral.

DD: OK. Do you think it's possible for someone to improve their pronunciation, either with a training like this or in any other way?

CYG01: Yes.

DD: Yes. Which way do you think would be the best way?

CYG01: To have someone watch movies all day.

DD: To have them watching movies all day. Like you did.

CYG01: Exactly.

DD: But did you pay attention to improve it?

CYG01: The movies that I used to watch were… basically… at the beginning…

DD: Do you think that they helped you… let me put it in another way… that at the beginning your pronunciation was not very good and then it improved when you were watching movies?

CYG01: I thought so…? I don't know…

DD: So you think that your pronunciation would be worse if you didn't watch movies.

CYG01: I think so, yes.

DD: OK. And that they helped you.

CYG01: Yes.

DD: OK. So you think that it is possible…

CYG01: Me, personally, yes. But if someone just watches the movie for the… just for the sake of it and does not pay attention when they speak…

DD: So it also needs some conscious effort.

CYG01: Yes.

DD: Did you do that? Were you paying attention?

CYG01: Yes.

DD: OK. Good. Very good. So… if, after the training we did, I was to tell you that you improved, would you have more motivation to try even harder? Or would you not care, it wouldn't make a difference to you?

CYG01: Hmm... I take a neutral stance again.

DD: Neutral. Do you think that you have motivation to improve your pronunciation or are you…?

CYG01: On a scale up to 10… a 6.

DD: 6. OK. So your motivation is 6 on a scale to 10. So, if I get this right, if it comes with little conscious effort you want it, but without making anything particularly…

CYG01: Yes.

DD: OK, I see. So… other than the training we did, would you do anything else to improve your pronunciation? After… and even after that training, I mean, were you paying more attention to videos or to your teachers or… or neutral…

CYG01: No.

DD: No. What about doing anything in the future?

CYG01: As I said... without...

DD: …much effort.

CYG01: ...much effort.

DD: OK. As long as it's easy.

CYG01: Yes.

DD: OK. Do you think that your pronunciation was improved in any way with the training that we did? I mean in vowels, in words, in sentences, in your overall pronunciation? Do you think that there was a change?

CYG01: On those days when the training took place, yes. For the next months, if you don't practice, no.

DD: So you had some change on those days but not…

CYG01: Yes. If you use it, if you don't speak English…

DD: OK. So… Do you think that this training can help the… those learning English?

CYG01: Yes.

DD: Yes. In what way? I mean, can it help with their pronunciation? With what? In… in understanding that there are some vowels that may be more difficult?

CYG01: It can help in general… In vowels, in knowing to… to… listen correctly to what the other person wants to say… irrespective of whether they will understand some and they will not understand others, if they speak to English people. But generally speaking, yes.

DD: That it helps. And it’s something that we can incorporate into schools, private lessons, with adults, with children… With whom do you think it could work, with everyone, with some people?

CYG01: Mostly with children.

DD: Mostly with children. In the groups of… let’s say in our private lessons?

CYG01: Yes.

DD: OK. With adults, you don’t think that… If an adult is learning English now, in an adult group, and he is told to do this training as part of the lesson…

CYG01: It will help. But, it will help more if they are just now learning the language.

DD: For those that… during language development.

CYG01: Yes.

DD: OK. Good. We are done with the questions. Any other comments?

CYG01: No.

DD: OK, very well, thank you! Let me stop the recording.

E.2 Interview with CYG03 (original in English)

DD: Hello!

CYG03: Hello!

DD: Just to let you know before we start that this is being recorded. Is that OK with you?

CYG03: Yes, it's totally fine.

DD: Thank you very much. So… Ohh uh… in this interview, I would like to get a little bit more information about your use of English during your school years and university years. Similarly to the questionnaire that you've responded to before you started your participation. So in your questionnaire, you state that you do have relatives who live in English speaking countries. So could you maybe tell me a little bit more about your relationship with those relatives who live in English speaking countries? For example, how often do they visit you? How often do you visit them? How much time do you spend together, whether you speak English or Greek when you're together, etcetera.

CYG03: So it’s been years since I started visiting my relatives in England and every time I try to use the English language, and then I started talking with them only in English and, however sometimes we're talking in Greek when it was more difficult to me to express myself in English. Uh, and I'm visiting my relatives… three or four times per year.

DD: Approximately for how long?

CYG03: For one month or two weeks. Something like that. Yeah.

DD: In total? Or each time?

CYG03: Each time.

DD: OK. And you would say that you speak more English with them?

CYG03: Yes. Because we have foreigners with us and they don't understand Greek, the Greek language. So we have to speak in English.

DD: Okay. And are they Greek-Cypriots who live in England or are they native speakers of English?

CYG03: No, they are Greek-Cypriots and they are living in England for for many years, yes. Five years and above.

DD: OK. And do they also visit you in Cyprus?

CYG03: Sometimes, yes.

DD: Would you say approximately once a year?

CYG03: Ohh no. I think once in a 2-year… once every two years.

DD: So for about how long?

CYG03: Two or three weeks, it depends.

DD: When they are in Cyprus, do you prefer to use Greek or English?

CYG03: No, we still talking mainly in language… in the English language.

DD: OK, very well… OK, now I'm gonna take you back to your school years if you remember. So do you remember, my first question is, do you remember any occasions where you spoke English with any friends or classmates or acquaintances, whether you were at home, in social life or at school?

CYG03: Yes, I did. Many times. I actually, I really like the English language. So I was not shy to speak in English. And in the classrooms during the activities that we had in the English lessons or sometimes with my friends, because we had many students from Erasmus, we were speaking in English.

DD: Okay… So it was it mostly in the English classroom or outside it as well?

CYG03: Both of them, both…

DD: So maybe with… um… the people in your social life?

CYG03: In, in well, actually the people in my social life, it's… they are my relatives. I however, however, I have friends from England, so I'm still in the English, still using the English language to communicate.

DD: Do you mean now? Or when you…

CYG03: Now and when I was at school, OK. And with my teacher, my English teacher, I was talking in English.

DD: You mean outside the classroom?

CYG03: Yes. And outside the classroom and in the classroom because she was very excited with me, my performance during the class time, yes.

DD: Okay so… I'll take you back to your English classroom. Hmm. Um, do you… Can you give a brief description of your experience in learning English at school? So, um… think about your teachers, the language of instruction, whether the teacher used English or Greek, whether you had any activities, what the syllabus included. Like, did you have speaking activities, listening activities…?

CYG03: Shall I, shall I include the private lessons? Yes. So I don't think so that the school teachers were very experienced… not experienced…, very… professionals actually I can say or in on their teaching method. We were using more the Greek language to explain the things in English and we were only using the English language in the writing activities in the book. In the classroom, we were talking in Greek mostly. Yes, yes. But in the private lessons, I can say that I learned from there. We were only talking in English and students that had a problem to express themselves and they were talking in Greek, the teacher was answering and providing them correctly the answer in English. So it was it was making you feel comfort to talk in English and to not be afraid to make mistakes.

DD: Yes. And were you comfortable in using English?

CYG03: Yes, I don't have problem. Yes. And then. Yes.

DD: Yes, very well. And… oh, in addition to the English language classroom, to what extent did you have any other opportunities to use, either speak or listen to English, for example, were you watching movies or were you watching lots of YouTube videos? Did you have any other opportunities where English was used?

CYG03: Yes, I think when I was watching movies. Yes, I like to watch movies with the UK accent, English accent and the UK or the US...

DD: The UK or the US accent?

CYG03: The UK, I prefer the UK accent and the way that they are talking in general… and on the YouTube I didn't prefer it a lot because they were using several accents and it was quite difficult for me when I was learning English to talk with their accent, yes.

DD: Very well, so… When you say you liked to watch movies, how often would you say you did that? Like, once a week, once a day…

CYG03: I can say once a day when I didn't have a lot of study, so frequently. Yeah.

DD: OK. Very well. And during all of these experiences you've described during your school years, can you recall to what extent you used in English with native and non-native speakers? So, was it mostly with non-native speakers? Was it mostly with native speakers? Was it balanced?

CYG03: It was mostly with native speakers due to my visits in England.

DD: OK, what about your teachers at school?

CYG03: The classmates and my teachers were not native speakers… they were Cypriots. However, I can say that we were communicating often with them in the English language, but not with the way that I was communicating with the native ones.

DD: OK, so you did have… ohh your teachers were non-native speakers… exclusively or mostly?

CYG03: Yes, exclusively, exclusively in school and in private lessons.

DD: Yes, OK, but you're… with your relatives, we are talking about people who have been living in the UK for a long time, OK. Now let me move on to your university years now, similar questions, but for more recent periods now, do you still have any English lessons at university?

CYG03: English language lessons yes, all my lessons are in English.

DD: No, I mean English lessons, learning English.

CYG03: No, no, I don't. No…

DD: Did you in previous years?

CYG03: …but I had in my previous… Yes.

DD: OK. OK. So how were those compared to your school years? Like, did you use English more at university during those lessons of English language…

CYG03: We were speaking only English in the university level, hmm… um… I cannot compare with the high school, my school experience with the university one. It was totally different.

DD: In what ways? Like, was the syllabus different…

CYG03: And the syllabus and we didn't have the Greek… language, at university.

DD: What about the teacher's language? I mean, were they native speakers, non-native speakers…

CYG03: They were non-native speakers. However, she was… because it was a woman. She was very familiar with the English, the language, so yeah…

DD: OK. So, you would say that she was more native-like…

CYG03: yes, we… yes.

DD: OK, but not a native speaker. OK, good. Do you have any comments about your experience at the university as concerns the use of English in the classroom, in all of your classrooms? So, in any other modules that you have, do you find it difficult that it's in English? Is it mostly in English? Would you prefer to participate in the classroom if it was in Greek? Do you avoid participating perhaps because you have to speak English and you don't feel comfortable doing that? How do you find this experience of… having to speak English in the classrooms at university?

CYG03: Actually on the part of understanding, I can understand everything that they are saying, the class. However, sometimes I don't feel comfortable to express myself in English and due from our foreigners that we have that they are from Erasmus… not because I'm trying to speak in English… but I don't know. Sometimes I cannot express myself with the way that I want, and if I was talking in Greek, I was gonna express myself much more differently than in English. So sometimes I prefer to not participate in the class and just listen to my lecture.

DD: Yeah. And would you say…

CYG03: On the other hand…

DD: Go on…

CYG03: …on the more… actually on the lectures on data, when it's more… economics and accounting basis which are more numerical things, I'm not expressing myself, but on the more theoretical lectures, I can express myself in… on the on the, on the one hand, on the… partly… Actually, I can partly express myself. But I keep avoiding speaking English.

DD: So the difference between the university and your school years is actually your classmates.

CYG03: Yeah. I can say yes.

DD: Now, they… the classmates, the diversity makes you more self-conscious perhaps… And you got shyer than before?

CYG03: Yeah. You can say that, yeah.

DD: OK, alright. OK. Ohh, during the university years, to what extent do you speak English at the university with friends, with classmates, with acquaintances, at home, in social life? So, in addition to your classroom, your lectures, how often do you… would you say you speak English?

CYG03: Very frequently… speaking English like, with the people that are from other countries, the foreigners. I still communicate with them, but not the things, not the part from our lectures, I mean actually socially we do communicate. Yes. I don't have problem with that.

DD: OK. So would you… do you have any classmates where you hang out with… with them outside the classroom, like…

CYG03: Yes…

DD: …socially? So, you talk to them in English.

CYG03: Yeah. I would say that happens often. Yes.

DD: OK how often… Approximately in a week?

CYG03: Well, I have lectures three times per week and two times of them I'm hanging out with them, with them.

DD: OK, good. Very well and… now moving on to the question about other opportunities to listen to English, like watching movies and YouTube videos. Do you still do that… at university?

CYG03: Yes, I do. But not that frequently as I was doing on my high school years.

DD: Okay. So… like, two times a week?

CYG03: Ohh no, I cannot say that. I can say two with three times per month.

DD: Okay so much less frequently than before.

CYG03: Yes.

DD: Any other opportunities… that came up later? So in addition to watching movies, maybe you're listening to podcasts?

CYG03: Yes, I was just gonna say that. Actually, I do like to listen to several podcasts on different topics and also, um… videos of how to improve yourself on self-improving videos, which is again in English.

DD: And do you do that often?

CYG03: Yes, quite often actually. When I have free time I can watch those videos.

DD: OK, hours approximately per day?

CYG03: Maybe I can say 30 minutes, 30 minutes per day. Yeah.

DD: And are they… do they include people that are native speakers of English?

CYG03: Yes, native speakers.

DD: Can you recognise the variety like, is it American English? Is it…

CYG03: Ohh yeah, it's pretty American… actually… from different nationalities, people that are talking in English with the both accents and American and English and the British, British…

DD: OK. Um, what else? Now, out of all of these that you we are talking about your university experience. Would you say… to, to what extent would you say you're using English with native speakers and non-native speakers? So is it balanced… Do you use it more with native speakers, more with non-native speakers?

CYG03: I think more with native speakers because I prefer… and due to my university, and I prefer to speak my native language with the people that I know the native speakers.

DD: You mean the language… not, not the English language…

CYG03: The English language, and I prefer to speak them with the native speakers. But the… then I'm trying not to use that much the English language. So I don't forget my own.

DD: OK, yes. But do you have any additional opportunities, like are your teachers…

CYG03: No, I don't think so. No, I don't.

DD: …or is it the same as with your school years, the amount of contact you have with native speakers?

CYG03: I think it's the same amount with my…

DD: So it's just your relatives and some friends…

CYG03: Yes.

DD: …you used to have when you were at school as well.

CYG03: Yes.

DD: OK. Nothing in addition to that… No… OK, good now. Ohh. Now when you… during your education in general, did you have any teachers or tutors at any point that paid attention to the student pronunciation? And if yes, can you give me some examples of what they did? For example, did they make corrections? Did they teach you specific sounds, did they include any listening tasks asking you to pay attention to pronunciation? Did they normally ask you to watch videos or movies? Did they do anything to actually…

CYG03: No, I think they didn't have problems with the pronunciation. Especially the teachers in my high school. However, the teachers in my private lessons, I cannot say that they were engaging… or encouraging us to talk with the really… or anyway with an accent. But they were really surprised, let's say when I first visited England and I came back, I had a different pronunciation… than before visiting the country. So they were very excited with my pronunciation. They were encouraging me to keep using this accent so I don't lose it… Yeah, but no specific. But no, no, no, no, not at all. Because they were not encouraging the other students.

DD: OK, so no specific activities.

CYG03: No, no.

DD: Okay. Now, what is your own opinion… your personal opinion about having a good pronunciation in English? Is there… do you think… is there any way that your English pronunciation may, might affect your current or future personal or professional life?

CYG03: I think, um, actually I don't have… Um… Both. Um…

DD: Ohh you can express yourself in Greek if you…

CYG03: No, no, I will find it. And I think that… When you're using it, when you are talking another language which which is not your mother language, it's the best thing is to use the pronunciation that is expressing yourself, like to don't try to copy, let's say the British pronunciation because we are not British and it's quite funny when you're trying to talk in a different accent. However, if you are trying to use these accents trying but with your best… way, I mean don't overdo it.

DD: Okay. So you… Um, what do you think? Um… what is your… So your opinion about having a good pronunciation? Would you… I don't want to direct your attention, but if I understood it correctly, a good pronunciation… Would you define it as a good pronunciation for each individual?

CYG03: Yes.

DD: And you should use that irrespective of whether it's a UK accent or your own accent.

CYG03: Yes, like, don't try to adopt the British native accent if you are, let's say, Greek, Bulgarian, Russian, because you will not. And you will not talk it as they are talking as well.

DD: So it's not feasible.

CYG03: Yes, I don't think… but not to everyone, but not not to everyone... Yes, maybe.

DD: OK, good. Do you think your… your own English pronunciation affects your professional or personal life now or in the future?

CYG03: I think yes.

DD: In which way would you say?

CYG03: Because… Um, I have… I'm in touch… In many ways, with the English language. I do. I have a an accent and the people that can hear me talk in English they are very surprised. And my first employer, she was very surprised and she preferred… Um, to keep me in touch with our English customers, let's say when they were coming in. Because of my accent.

DD: So, in… let's say in the future, you're talking about jobs. So professional life. Would you say in the future, it can have a positive or a negative effect…

CYG03: I think it shows that it will have a positive effect if you have a pronunciation, yes.

DD: Perfect. Now have you on your own… Have you ever tried to practise your pronunciation skills many times? You know, how did you do that?

CYG03: Um.. actually, I was listening to different native speakers how they pronounce several words, and I was trying to pronounce them with the same way in front of my mirror.

DD: Any other ways?

CYG03: Ohh, I'm talking with native speakers and yeah, only those two ways actually.

DD: Very well, OK. Some people believe that it is more important for people to understand what they say, even if they have a foreign accent and others believe that sounding native-like is equally important. What is your opinion about this disagreement?

CYG03: Um. You mean to understand people…

DD: To understand what they are saying even when they are talking with a foreign accent. So some people say it's more important for people to understand me no matter how I speak, no

matter if I have an accent, a foreign accent. And some people say that it's very important to have a native-like accent.

CYG03: No, I don't think so to have a native accent it is important. OK, so I think you're trying you... Actually, you can try to be understandable as much as you can. So you can express yourself to the others.

DD: To be understandable.

CYG03: Yes, you're… exactly. Yes.

DD: Would you say that it's important to do your best, though?

CYG03: Yes. I mean to trying to be understandable as much as you can, you can. Yes, exactly.

DD: Okay, I think you've answered this, but let me ask you again. Are you confident in speaking English? And if not, is this because of your pronunciation?

CYG03: No. I'm very comfortable to speak in English as it looks. I know that I'm making mistakes. I know that I'm making mistakes, but I don't have problem with that.

DD: No worries, okay. Okay similar uh question, do you ever feel uncomfortable speaking English with native English speakers, other Greek Cypriots or any other non-native speakers of English? So three groups, yeah. Greek Cypriots, native speakers of English and other non-native speakers of English.

CYG03: With the non-native speakers in English and the Greek Cypriots, I don't have a problem at all. And on with the native speakers, again I don't have problem. And when I'm talking with them, I'm asking them to correct me when I'm making mistakes while I'm talking because with this way you learn much more, I think, because we're practising the language and the same time you can understand the mistakes that you are doing while you are… Yeah… Doing them. Yes. While you're speaking.

DD: You mentioned before at the university, you are not very comfortable expressing yourself in class.

CYG03: Yes.

DD: And that is because of your classmates?

CYG03: Yes.

DD: Is it because they are…you feel that those particular people are judgmental, for example?

CYG03: Yes.

DD: OK, irrespective of whether they are native speakers of English or…

CYG03: No, they're not native speakers in English. But I can I feel… something like they gonna judge me, OK? Not not…

DD: The… because this question is similar, not because they're non-native speakers, but because these specific individuals…

CYG03: Yes.

DD: …make you feel like…

CYG03: Yeah.

DD: OK, alright then. Now do you ever feel worried about your pronunciation? For example, are you worried that people might not be able to understand you or that it can cost you a good career opportunity?

CYG03: No, I don't think so. That they're going to cost me an opportunity… Um… career opportunity, but sometimes when I'm overdoing it with my accent, I think that people cannot understand me because it's what I said before when you're trying to overdo an English accent, to adopt another accent, you sound a little bit funny.

DD: OK, very well. Would you say that your pronunciation level is satisfactory and can therefore meet the expectation, the expectations of employers?

CYG03: I think yes, yes.

DD: Very well. OK, is there any particular future goal that you have which may require a good level of English, for example, do you think you might be studying abroad in a few years or you may want to get a job requiring high communication skills in English?

CYG03: I really do wanna study abroad. And I really want to improve my English while I'm talking, to talk more fluently, I mean.

DD: Would you say that having a good level of English would help you achieve those goals?

CYG03: Yes. Actually, when you are in touch everyday with the English language and this is not only with the English language, with any other language, when you are in touch with this language, you are learning more easily than using it once or two times per week and…

DD: You would say that it opens up opportunities for you travelling...

CYG03: Not only for travelling, and for jobs and for your future and so many things.

DD: And this is part of your plan.

CYG03: Yes.

DD: OK. Um, now. Why did you want to participate in this study? What… What were your expectations about it? And do you think it can help you improve your pronunciation? Do you believe this is possible?

CYG03: Yes, of course it is possible. And also… And it's a very good opportunity, I think for putting it on… to also mention and on your CV and to have it as an experience generally.

DD: Just to remind you that the training that you completed was optional. So why did you choose to do the training? Did you think that it might help you improve? Did you do it because you are curious…

CYG03: Yes, and because I was curious, but because I thought that it’s going to help me. And it was a good opportunity that time to take the training.

DD: Very well. If you, um… I don't know if you personally noticed an improvement in your pronunciation, but if you were to notice some of those improvements in your pronunciation after some training, would you feel more motivated to try and improve your pronunciation even more if you saw that it was working?

CYG03: Yes, of course. Why not? I mean, I saw that some things that we are pronouncing differently with the training, I heard them different, with a different aspect, so yes.

DD: So you would say that the training was effective for you.

CYG03: Yes, of course.

DD: Very well. In general, would you like to improve your pronunciation, or is it… Is this something that doesn't really concern you? Do you… Would you like to improve even further? Is it something that, if it comes naturally, then it can come, or how do you feel about improving…

CYG03: The process in mind is to improve my current situation because I'm still not talking with the same way that as I want, and I think it was not… If it was an opportunity to improve my pronunciation, I would take this opportunity very well.

DD: Um… Okay thank you, and do you think that the training helped you improve your pronunciation in any way in vowels, in words, in sentences? You mentioned that you felt like you had…

CYG03: In words, mostly in words. Yes. So in, in noticing maybe differences they… when they were talking to me. Yes, I could understand more, better than before the training. Yeah.

DD: Does it does it still happen to this day?

CYG03: Yes.

DD: After a year?

CYG03: Yes, because everybody... Everyone talks with another… with another accent.

DD: OK, very good. In general, this type of training that you've completed, do you think it can be beneficial for learners of English of any age? If yes, in what ways do you think… it's something that we can incorporate in schools, in afternoon institutes, in adult programmes, in individual training, individual learning?

CYG03: I think it can be beneficial in ages like high school 15 to 18? Um, they will help them with their way to listen to others in the listening part. It will improve very much uh… Their skills on listening.

DD: Why are you focusing on those groups? Why not younger or something?

CYG03: Because the youngers are not taking that much serious, I think, the language as the older ones, because the older ones they're gonna finish the high school and then they're gonna move on to their... Universities. So they're gonna want to study abroad. It will help them very much. Or if they want to go to university, which the basic language is English again, it will help them.

DD: So they are more aware of their goal.

CYG03: Yes, I think yes, not everyone, but mostly.

DD: Most of the children might not pay attention.

CYG03: Exactly. Yes. Because they're younger, younger children under the age of 15, they're using their private lessons. So will not have that much… and I think neither the teachers, will give that matter much attention to them.

DD: OK, very well. What, what about adult programmes? So would it not be suitable for them?

CYG03: With all adults that they are interested on the English language, yes.

DD: Motivated.

CYG03: Yeah. Yes.

DD: Okay very well. Any other comments you would like to add before we finish this interview?

CYG03: No, I think I'm okay. Thank you very much.

DD: Thank you. For your time. Um, I'll just turn off the recording now.

E.3 Interview with CYG08 (translated from CYG)

DD: OK, we are ready. So… You are participant CYG08. In this interview, I will basically ask you questions similar to those of the questionnaire, so that I can get more information for analysis. As I mentioned, this is recorded, do you agree with this?

CYG08: Yes.

DD: Ok good. Let's begin. I would firstly like to ask you about your relatives. You mentioned in your questionnaire that you have relatives who live in English-speaking countries. Can you tell me more about how often they visit you, or how often you go, how much time you spend together each time, if you speak Greek or English.

CYG08: Well, my uncle, my aunt's husband, he is Cypriot, but he lived many years in England, so in general, he feels more comfortable speaking in English, but we speak 50-50, maybe more in Greek, or a mix between the two, something like that. He lives in Cyprus permanently. The remaining relatives are from the USA and we have relatives in Australia too, who come during the summers, we meet with the cousins, with whom we speak English. We use Greek a little, we mostly speak in English. There were some summers when they didn't come, but usually, they come every summer.

DD: Hm... Approximately how many days a year do you spend together?

CYG08: Em... One week at most?

DD: OK so do you spend almost every day together when they come?

CYG08: Em… Because there are a lot of cousins, it may happen that four members of the family come, or we may have one or two of them, so it depends. But not always, they may not come at all in a year.

DD: OK.

CYG08: The other relatives in the USA… For example, my aunt from the USA has been in Cyprus for 3 years now, so I see her more often, but again, we speak a mix of the two languages, sometimes… A little Cypriot-Greek, a little English, it depends.

DD: OK, and do you go to your relatives abroad at all?

CYG08: I haven't been there up to now to be honest… neither to the USA nor… No I haven't been there ever…

DD: OK, in any other English-speaking countries?

CYG08: Generally speaking, with Erasmus programmes, we speak more… I have participated in some, and we speak in… in English… And at the University in general, most of my lectures were in English and we used English with our lecturers etc…

DD: No, I mean if you have visited… if have travelled to an English-speaking country.

CYG08: Oh… Where English is the first language. England…

DD: Yes, for example England, the USA, Australia…

CYG08: No, no, not to these.

DD: OK I will ask you about the University, to give me more details, but first I will take you back to your school years, as with the questionnaire if you remember at all. I wanted to ask about your school years. Do you remember any occasions where you spoke English with either your friends, or with classmates, with acquaintances or with people you didn't know, in your social life in general and at school, in the classroom let's say?

CYG08: Emm… I didn't have any classmates who spoke in English, and in our English lessons in general… we didn't speak in English… paradoxically… let me think about it… generally speaking, outside the classroom, if I happened to speak anywhere… I mean in more touristic environments etc., but I felt quite comfortable speaking English. I started having private afternoon lessons in the 2nd grade of primary school… I was comfortable with languages in general… I don't remember if you asked anything else.

DD: No, no, just to clarify, basically when you needed to speak English, it was in cases where for example you went to a different town, in a more touristic place, and there was a foreigner and you happened to speak English…

CYG08: Yes. Even though…

DD: …not more frequently than that.

CYG08: No, no. Even though, due to this, I didn't use to feel that comfortable speaking it before, whereas as I got older and as I practiced more, I felt more comfortable using it.

DD: Yes. Good. OK. Now, in school, either in private lessons or in class at school, can you tell me a bit about the experience you had when you were learning English? That is, your teachers, I mean what languages did they speak, what languages did you use in class, the syllabus you covered, if it included grammar, if it included speaking, writing etc. Whatever you can remember.

CYG08: At school, in general, the things we did were very specific. They spoke to us… I think only… In middle school I remember for sure that they spoke to us in… the teachers… the teacher in middle school was not good at all either in middle school… I don't remember if I had her for two years… In middle school in general, the English teacher was not good at all… And we focused mostly on grammar, on very basic things, and very basic things on vocabulary enrichment.

DD: What language did you mainly use?

CYG08: She explained in Greek. And then in high school, I remember, again it was very basic, again the same focus, but I think the teacher used to speak in English, but it was only for about one class period a week, or something like that… it was very little time… especially, for example, in the 2nd-3rd grade of high school when we had our major subjects etc., nobody paid attention to languages. Most of the work was done in the private lessons, where I had started from the 2nd grade of middle school… of primary school, and in the 3rd grade of middle school I got the IGCSE so I finished the lessons. And that summer I worked at the institute as an assistant for the kids who did their English homework there at the institute after they finished the lessons. That's all..

DD: OK, and at the institute, again, did you have grammar… What did the syllabus include?

CYG08: At the institute the teacher spoke to us… in the younger-level classes and up to the final year in Greek, and he/she focused more on grammar… rather than vocabulary. I mean, I still feel that my vocabulary is not that rich. And speaking… it was in the final stages that we focused more on speaking and listening.

DD: OK, and did you speak in Greek almost always, if I understood correctly?

CYG08: Yes, only in the final 1-2 years when the exam was on a more high level did we speak more in English.

DD: OK.

CYG08: But we still spoke in Greek, Cypriot-Greek as well.

DD: OK. And did you do any activities at all on… you mentioned speaking, listening?

CYG08: Yes, but very few, only in the final 2 years did we practice more, but I remember that listening was the most difficult part for me.

DD: OK, and I would like to also ask you in your own life, your personal life, what other opportunities did you have, where you used, either spoke or listened to English? That is, did you use to watch movies or Youtube videos, or listen to speeches in English? I mean on your own, in your afternoon.

CYG08: I used to watch, I can't say all the time, but I always had Greek subtitles, I struggled when I had English subtitles… this has changed in later years, but… I mean at 18 years old… And outside of this I didn't have any friends with whom we spoke in English… no, I didn't have any apart from…

DD: So generally speaking… I'm sorry I interrupted you.

CYG08: No, I was just saying, apart from the cousins, the relatives, I didn't with any friends.

DD: OK, so in general, basically during your school years you didn't have much contact with English, in the sense that you listened to it often or spoke in it often?

CYG08: Emm.. No, I think mostly at the institute… and from some movies and the like, and songs, because I liked English songs more than Greek songs… But I think to a moderate level.

DD: OK. Do you have… let's say when you were watching movies or any other video, did you have any preferences… or do you remember at all if they were in the American accent, the English accent… if you can tell them apart and if you paid any attention. Did you have a preference or…?

CYG08: I remember that the American accent came to me more naturally… and my aunt, because she studied in the USA… in general I heard her, with my uncle, who… who grew up in England, somewhat… the difference between my uncle… the British accent, and my aunt, the American accent. This is what I remember the most… and the British accent always seemed a bit strange to me… I used to watch some movies a little bit but I think most of them were American, the ones that I watched.

DD: American… Do you have a personal preference, an accent that you like more?

CYG08: Emm… Truth is, I like the American accent more, I feel that it's more… accessible, more… of the younger generation kind of… but the British accent is also interesting, I like that

it brings out a quality, a… culture, kind of… so both of them, I kind of like to play with accents sometimes…

DD: OK. So my next question was, if you remember at all, in your experience at school, if you interacted more with native or non-native speakers, but I think you covered me with what you said. The only think I'll ask you on this matter is if you had any… were your teachers at school or at the institute, at least some of them, native-speakers, English or American, or had English as their first language, or was the majority of them Cypriots?

CYG08: At school… it was the Cypriot accent type. At the institute, he was a Cypriot who had been in England for many years, who was… but I can't say that it was that British accent, that intense one, it was somewhat more mild. But definitely not the American accent… yes, that.

DD: OK. So, moving on to more recent years now, to your university years. At university… I don't remember exactly, are your lectures in English now?

CYG08: I'm done with the lectures now… Most of them… most of them were in English, I only chose very few of them in Greek… That is, about 85% were in English and the other 15% in Greek, let's say… and with the lecturers we mostly spoke in English, but if anything was needed, let's say in Greek, they would reply to us there was no problem. But in general, they preferred speaking in English.

DD: OK. And apart from your lectures that were in English, did you also have any English lessons, that is, where they were teaching you grammar, vocabulary etc. to improve your English?

CYG08: We had a module that was like that, for English as a language, and then we had another module that was about English but in relation to our studies, e.g. how to write a CV… various letters, etc., emails, stuff like that, it was about technical writing. And the other one was for communication, that is, how to make a presentation, which was again in English, and we did presentations that were in English… on various subjects.

DD: OK. Do you remember at all… in these modules, that were mostly about English, do you remember at all how were your lecturers… if they were Cypriots, what languages did they speak, what did the syllabus include?

CYG08: In the first two modules, we had a teacher who was from… what country was it… not Ukraine, not Latvia…

DD: From another country let's say, not from an English-speaking country.

CYG08: No, but she was in England for many years and her accent was really good… and in the other module, she was English, I think? Yes. Yes. So we spoke only in English. The second one, her accent was fully… very clear, good English, let's say.

DD: If you remember or if you can tell them apart, was it American, British…?

CYG08: It wasn't that British… that very… London accent let's say, but it was… I can't say it was American.

DD: OK.

CYG08: I think at the university they are trying to… the accent to make it… I don't know if she speaks with a purely British accent in other contexts, but they are trying to bring it down a little so that it's understood.

DD: Yes, I see. OK. As concerns your experience at the university, in relation to these modules of English, I understand you only spoke in English.

CYG08: Yes.

DD: Do you remember if this was difficult for you, did you prefer not to participate because it was in English? Did you prefer to speak in Greek? How did the fact that you spoke in English make you feel?

CYG08: During the first semester at the university I got the modules in Greek except for the one module that was English, and we had an exam at the beginning for English, because normally some… you did an exam, because I had completed the IGCSE long ago… so that they could check your level, and they put me in… in the second lesson directly… that is I skipped one… some others had an extra lesson before the one that I started… and… I remember that during the first semester I struggled with English, I didn't do well in that module… but, because I had come back from Greece, where for three years I was not exposed to English at all, that is, everything was in Greek, they don't have English anywhere, not even on the computers, so I lost all contact with English completely. So I struggled during the first semester, but then during the next one I chose them in English, and because at university, in the English classes there were mostly Erasmus students etc., so I also spoke English outside the classroom and this helped me a lot to… to get used to it more easily… and in general, in classes, I was the kind of person who preferred to speak it more, even if I struggled, so that I would start having more exposure to it and familiarise myself with it even if I felt like I didn't speak it as well or even if I struggled, so that I practiced.

DD: I see. OK. And you mentioned that you had a lot of Erasmus students… did you hang out… you mentioned that you used it outside the classroom as well… was this common, that is to speak English with friends, classmates, acquaintances outside the classroom? In your social life in general.

CYG08: I used to hang out a lot with a girl from Germany, who was studying in the Netherlands before and came through the Erasmus programme, who spoke it quite well… and with other folks who didn't speak English that well, and there were some other people… like socially in class… who were in the "global semester", from the USA, and we used to talk a little… and also, I used to work at the university, where everything was in English, both the emails we used to send and the communication, because some people only spoke English…so it was in my life from that point onwards, during the 3 years, it was about 50-50, that is I spoke Greek as much as English.

DD: Do you remember at all if it was with English-speaking people, that is people with English as their first language, or if it was with people from other countries, this experience that you had?

CYG08: From England, not so much I think.

DD: Or the USA, or Australia. I mean other than relatives.

CYG08: Emm… No, it was from other countries. That is Germany… from various others, yes. And from outside the European Union.

DD: OK. And the last thing for this section, during your university years, do you remember any additional things that you may have done where you had exposure to English, that is movies, videos etc.?

CYG08: The Erasmus programmes I joined, when I went abroad for seminars etc., and generally through those… and the applications I made etc… where I had more exposure.

DD: How long did you spend in Erasmus programmes, that is, how many times did you go, for how long…

CYG08: Up to now, I went to… 5 times abroad and they were… 2 of them were for one week and the others were for two weeks.

DD: OK. A lot of time abroad. OK. So you had a lot of experience with English, but with people from other countries, not from English-speaking countries.

CYG08: Yes, it was a minority, the people who… especially the British accent, almost none at all.

DD: Almost none at all. OK. Let's move on to your education in general. If you remember, were there any teachers at school, at any point, whether it was in middle school or high school or the university, did you have any teachers who paid particular attention to the students' pronunciation?

CYG08: At the institute, but again, in the last 2 years, not so much before then. I mean they focused more on grammar. At school OK, they focused somewhat on pronunciation but nothing remarkable.

DD: OK, and at the university, nothing at all?

CYG08: Basically at the university, because there were people from various countries, they wanted us to feel comfortable expressing ourselves and communicating so… but indirectly, let's say, they repeated things when we said something wrong.

DD: When you said something wrong, OK.

CYG08: That is, they would reply in the right way. But it wasn't like they would make us feel bad, that we said something wrong, so as to encourage us to participate.

DD: OK, yes I see. OK, and at the institute where you mentioned that they did something, can you give me a few examples? Did they correct you, did they teach specific sounds, did they do listening activities and asked you to pay attention to how they spoke, or did they play videos and movies all the time and told you to pay particular attention to anything?

CYG08: During the first years when I started at the institute, they focused a lot on the different sounds through the alphabet etc., and then as we learnt… and during the… there were CDs there and we could borrow them to watch them at home etc. and then, in the last 2 years, when we had more intensive speaking exams, we would form groups and we would practice our speaking, and we would stop at some points so that they could correct us or so that we would practice something on our pronunciation or on… on various aspects.

DD: OK and when you did those things, when they corrected you or you focused on pronunciation, did they do it using their own examples, did they repeat with their own pronunciation, or did they ask you to listen to examples from the internet or from CDs or from any other place so that you could listen exactly how an English person would say it?

CYG08: In their own pronunciation.

DD: In their own.

CYG08: Yes.

DD: OK.

CYG08: I think if there were other ways too, it would have been more helpful. I mean, like what we did in the experiment together, it was very helpful to listen to something by different voices.

DD: I'm glad. OK. So, now we are moving to your own opinion. I would like to ask your opinion about having a good pronunciation in English. Is there a goal you have that might be affected by this? I mean do you have any personal or professional goals that might be affected by the way you speak English, by the accent you have?

CYG08: For me, at this stage, the most important thing is to communicate, to be understood and to understand the other person and to be able to communicate by expressing what you want to say, I mean, which is much more difficult in a language other than your native language, especially with some expressions. But personally, ideally, I would like to have a nice accent. I don't know, I like a nice accent both aurally and aesthetically… So I would like that, but with this criterion, because, in psychology as well, when you have to do with clients, patients etc. it's important to speak clearly and to achieve comprehension and to you have some stability.

DD: Exactly. OK. So it's important to the extent that communication is achieved. Not necessarily to the extent that you sound like a native speaker.

CYG08: Ideally I would like that, but I think… it's impossible to… unless you live abroad for years, to reach that level, but vocabulary or grammar definitely come before pronunciation. I think with pronunciation, it's when you reach the point where you master the language.

DD: It's the finishing…

CYG08: Yes. Which is the ideal, but I'm not sure how feasible it is to reach that level.

DD: OK. I have a question about this later on. So… I think you've answered it already but we'll move on and we'll see. So, apart from the training we did, have you ever done anything else… you say that personally, you would like it… Have you ever practiced in any other way, either alone in front of a mirror or by watching movies and repeating things… have you ever done anything to improve your own pronunciation?

CYG08: I used to do that, when I was younger, I would practice in front of a mirror, I don't know, on my own. When I saw anything for example on TV I would practice it, I think even before I started the private lessons, I mean I was the one who asked to start the lessons earlier, my parents wanted me to go later. Because I would see my siblings and I would see them doing their homework and I wanted to learn too. And then, with movies, I always… especially if I could pause them, I would pause on some words that I didn't know so that I could look them up and practice them a little bit on how they were pronounced. Because I'm an audio-visual learner as well, so the combination helps me in general. But I don't know if I could do more on this, I mean, I don't know if it was adequate in terms of the pronunciation component or if I could have done something more. That's why I liked having this opportunity to participate in this research, in that I could learn a few more things.

DD: OK, so the next thing… you've partly answered that already, if you want to add anything. Some people believe that it is important just for the other person to understand you even if you have a foreign accent, and others believe that it's equally important to be able to communicate and to have a good English accent, that is to have an accent that is as native-like

as possible. Which position do you support more, what is your opinion? Is it that it's equally important to have a good accent and to be understandable, or that it's more important for people to understand what you want to say irrespective of whether you have a foreign accent?

CYG08: I used to think that pronunciation and sounding native were important before, and I used to think that stereotypes play a role as well, and that… for example based on stereotypes a client might think that "she is not native so she won't understand me, and therefore I prefer someone who's a native speaker" etc., but along the way, at the university, I realised that, the fact that I felt that I had to have a better accent was kind of an obstacle for me in terms of expressing myself as much as I wanted to, so the percentage of importance I assigned to having a high level of pronunciation somewhat changed… and I saw now that… through teachers for example, who had a really good vocabulary, grammar and they were very good, but they still had that Cypriot element in their accent even though they lived abroad for years… I saw through examples that it doesn't play such an important role and that… the stereotypical part, when you are good at your job, sounding native-like can be overcome in a way, so the percentage changed for me and I believe that… ideally, I would like to have it, but the percentage is not as high as it used to be.

DD: OK, you put it quite nicely. OK. I think you mentioned something about this next one too, but in case you'd like to add something. Are you… do you feel confident speaking English? Are you comfortable with speaking English or not?

CYG08: Emm… I feel more confident in the field of psychology, where I'm more familiar with the vocabulary and the expressions from my studies. There are aspects in everyday life, because I'm in Cyprus, and there is vocabulary I'm not familiar with in more everyday things, and it may be a little challenging for me and make me feel somewhat less confident, but it doesn't stop me from… learning. What I feel is… where I feel less confident is in things that I haven't processed yet and I have to express them on the spot, which I can do more directly in Greek because I'm still processing them and I haven't yet… in expressing them… I find it difficult to express things exactly as I have them in mind.

DD: OK. But it doesn't have anything to do with… whether you feel confident or not doesn't have to do with your accent for any reason. You are ok with your accent, you are comfortable, it doesn't bother you, it doesn't affect your confidence or anything.

CYG08: When there are people that have a really good accent, I feel somewhat… less confident. When there are people for example from other countries, who also don't have the best accent, it doesn't bother me at all. But when there are people that have a really nice accent, I don't know, it's somewhat like… like it's coming out… at the time we talk, it's like I mimic their accent a little, so I'm closer to theirs, so this makes me feel more confident afterwards. And I notice myself afterwards, when I leave, I speak differently.

DD: OK, this is good. So, my next question is if you feel… if you don't feel comfortable speaking English with… English people, native speakers, or if you don't feel comfortable speaking with other Cypriots, or if you don't feel comfortable speaking with people from other countries. You partly mentioned some things, is there a group of these 3 that makes you feel uncomfortable speaking English with?

CYG08: I find it a bit difficult when they have difficulties with English on their part as well, and there is this mutual thing where we try to say some stuff but we find it difficult to express ourselves. I mean, in that case I feel like I get stuck and… I try too to find the code. Because,

for example, when I went to Latvia through Erasmus, they didn't speak English well and they were trying to express themselves and I could feel their difficulty in speaking English and them closing off and I felt too that this blocked me a little bit in… in communication.

DD: Yes…

CYG08: But in the end, ok, with practice… they improve and you improve too so you get there… this, and with native speakers a little… I may feel a bit awkward, but along the way, with conversation, it's ok.

DD: OK. With Cypriots, not at all?

CYG08: It depends on the individual, if they feel comfortable, then I feel comfortable too. If I see that… they prefer speaking in Cypriot Greek and they feel uncomfortable speaking in English, I don't prefer it either, that is, it depends on the…

DD: On the individual…

CYG08: Yes…

DD: OK. Moving back to accent. I think I get the answer from what you mentioned already, but let's say it clearly. Do you ever feel worried about your accent? For example, do you worry that you may not be understood, that they may not understand you, that you may miss a job opportunity due to your accent… do you have a job in mind that might require impeccable English and you may miss the opportunity due to your accent? Is there any domain in which your pronunciation worries you?

CYG08: Emm… I was thinking that because I want to make some applications in England, which have to do with clinical psychology, so I don't know how important it is for them to speak with… with a good accent. But because I'm mostly interested in Scotland, which is more… multicultural, I feel that it won't be a problem, whereas I don't know if they would be stricter in London. Emm… and in general I was thinking that… I don't know in a PhD or anything if they would consider it much more important. But this is something for the future, so at the stage I am now, when I'm also looking for an internship in… the European Union, because they come from different countries, they are not concerned about having the perfect accent, so… it doesn't concern me that much.

DD: OK, so basically, you think that unless they want an English person, your pronunciation level is good enough to meet their expectations, of any employer or teacher or anyone… Do you believe that you are understandable?

CYG08: Emm… in my field, I believe yes.

DD: OK. You've already mentioned your future goals as well, so moving on. So, we are almost done. Why were you interested in participating in this study and what were you expecting, I mean did you expect that it can improve your accent, do you think that this is possible?

CYG08: Emm… yes, I wanted to… basically I saw it as an opportunity to… it was also during the period when I was starting my dissertation as well and I could see that the issue with participants is difficult to I participated in other people's studies as a mutual-assistance thing between us, and I liked this particular study because I felt that I would also learn something from it in relation to… because up to now, ok, I paid some attention to it, but I had never… this seemed interesting, to see the differences between the American and British accent more in practice… I feel that it helped at some level, but… I was thinking about it now that we were

going to have this meeting, that… how much do I remember? But I feel that it was more experiential than something that I can remember like… in memory… that is it was more… experiential, yes…

DD: Something of the moment you mean… like it helped you at that stage?

CYG08: I feel like I can't remember exactly, like rules 1, 2, 3, but I feel that if I see it again I will be able to remember it and use it, something like that.

DD: OK…

CYG08: The thing that made it difficult for me… or is it not part of the question?

DD: No, do tell.

CYG08: The thing that made it difficult for me was that I didn't have the feedback at the time to check if I had done it wrong… and I didn't know if I got it wrong or right… I don't know if that helped… it was more… maybe that's what helped a bit in making it more experiential.

DD: Yes…

CYG08: But it made it a bit difficult for me, maybe because that's how we are used to operate both in the English lessons and everywhere.

DD: Yes. OK… If at that stage, or later, if you noticed that some aspect of your pronunciation was indeed improved, if you noticed some improvement on yourself, do you think that you would be more… you would have more motivation to try to improve your pronunciation even more? Do you think that it could… that it would be possible and therefore you'd be more motivated to… to do something extra to improve even more?

CYG08: In terms of the British accent yes, yes.

DD: OK.

CYG08: Because it's a bit more difficult, it's more difficult for me, so if I noticed that it has a bigger effect, I would try harder.

DD: OK. You've answered the next one too, but in case you'd like to add something… would you like to improve your pronunciation, or is it something that really doesn't concern you at all?

CYG08: I would like that, yes, I would like that.

DD: You've fully answered this question already.

CYG08: Yes, it seems challenging to me basically… yes, more like a game, like… yes, it's interesting.

DD: OK. So… and personally, do you think that the training helped you improve either in vowels, or in words, or in sentences, or generally in your pronunciation? Did you notice, did you feel any change, any improvement?

CYG08: Emm… I remember that it helped me a lot on some points… of course since then, I haven't used it anywhere… I haven't used it anywhere in particular to be sure about how much it helped… I mean if we had a test again now, that could somewhat help me understand better if indeed…

DD: If you still remember or not.

CYG08: Yes… and I remember that some particular ones, which were the most difficult, I'm not sure if I could get them right again.

DD: I see, yes. OK. So, my last question, and then you can comment on anything if you want, is if you think that this type of training could be helpful for people who are learning English, in which ways do you think it could help, and if it's something that we can incorporate to schools, institutes, adult programmes, with children, with adults… do you think that it's something that could help, that would be good to have in our schools or with adults?

CYG08: I thought it was very interesting, because in a way I linked it as an experience with… the difference between formal and non-formal education, the fact that… in institutes, at schools, we got used to a more… particular way… not as… as a lesson… not so much through experience…but this puzzles me, that it would be very different to the way in which I was used to learning the language up to now, for me, so… this is what I was thinking, that in combination, if all the rest continues to… if learning continues in this way and we also have this, to what extent they could work in combination.

DD: Yes, I'm talking about a combination, to add this, let's say.

CYG08: Yes… this is what I don't know, in younger ages, because I did it at an older age, in a younger age I don't know if it would have confused me a little.

DD: OK, so you think it's better to have this in adult programmes.

CYG08: Emm… I think in younger ages too. It's just that in general, I think that the way… at least the way in which I learnt the language, it would be better if it was a bit different, I mean through more activities, through more… videos and other means, which in combination with this I think they would work much better. But, for example, this with the different voices and the same… this was very helpful.

DD: OK. Yes, I'm glad. So, that's all from me, do you have any other comments that you'd like to add, anything I haven't asked you?

CYG08: No, I'm good, thank you.

DD: Thank you for your time.