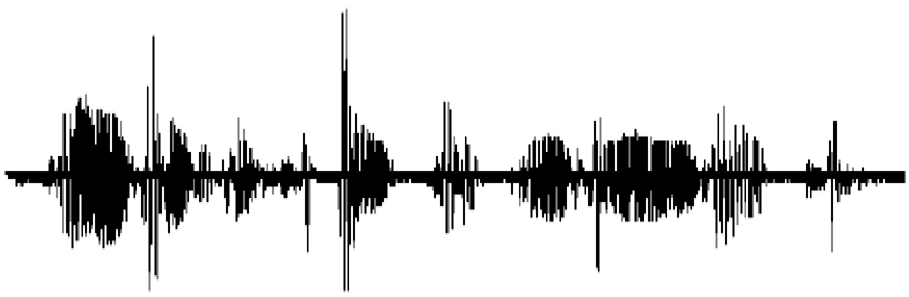
Plasticity in second language (L2) learning: perception of L2 phonemes by native Greek speakers of English.

A thesis submitted for the degree of
Doctor of Philosophy

By
Anastasia Giannakopoulou



Psychology Department, School of Social Sciences,

Brunel University

June 2012

Abstract

Understanding the process of language acquisition is a challenge that many researchers spanning different disciplines (e.g. linguistics, psychology, neuroscience) have grappled with for centuries. One which has in recent years attracted a lot of attention has been in the area of non-native phoneme acquisition. Speech sounds that contain multiple phonetic cues are often difficult for foreign-language learners, especially if certain cues are weighted differently in the foreign and native languages. Greek adult and child speakers of English were studied to determine which cues (duration or spectral) they were using to make discrimination and identification judgments for an English vowel contrast pair. To this end, two forms of identification and discrimination tasks were used: natural (unedited) stimuli and another 'modified' vowel duration stimuli which were edited so that there were no duration differences between the vowels. Results show the Greek speakers were particularly impaired when they were unable to use the duration cue as compared to the native English speakers. Similar results were also obtained in control experiments where there was no orthographic representation or where the stimuli were cross-spliced to modify the phonetic neighborhood.

Further experiments used high-variability training sessions to enhance vowel perception. Following training, performance improved for both Greek adult and child groups as revealed by post training tests. However the improvements were most pronounced for the child Greek speaker group. A further study examined the effect of different orthographic cues that might affect rhyme and homophony judgment. The results of that study showed that Greek speakers were in general more affected by orthography and regularity (particularly of the vowel) in making these judgments. This would suggest that Greek speakers were more sensitive to irrelevant orthographic cues, mirroring the results in the auditory modality where they focused on irrelevant acoustic cues. The results are discussed in terms of current theories of language acquisition, with particular reference to acquisition of non-native phonemes.

Declaration

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Anastasia Giannakopoulou

Acknowledgements

Besides the hard work, endless data sets and long days in the lab, the production of this thesis involves some valuable contributions from people who have accompanied me all the way throughout this exciting and intellectually stimulating journey.

I would like to thank my supervisors, Dr Maria Uther and Prof Taeko Wydell, for their exceptional assistance and guidance toward the completion of this thesis. They have been an indispensable part of my academic progress and intellectual formation. Maria has given me a deeper insight of academic writing and research. She was always there for me, encouraging me every step of the way, ready to help and answer my queries, even spending long hours with me in the lab, a wonderful teacher and mentor. I do not think I have the words to thank her enough for being by my side during this exciting, yet laborious, time! Taeko has been ever so helpful, always pointing me to the right direction, attending to my weaknesses, encouraging me every step of the way. Both, Maria and Taeko have been an inspiration and I thank them for believing in me. I would also like to thank Prof Anne Wichmann (UCLAN) who was my tutor in Applied Phonetics and Phonology in my undergraduate years and was in fact the one who inspired the initial idea for this current research!

This thesis was supported by funds from the School of Social Sciences, Brunel University, UK, that awarded me the Research Studentship stipend. I am forever grateful that it was made possible through this studentship to accomplish my dream for doctoral research. I would like especially to thank the Psychology Department staff and admin team for being supportive one way or another. Extended thanks to the technicians for their assistance on numerous occasions and for always responding to my requests for technical support.

I am very grateful to Dr. Reiko Akahane-Yamada of ATR laboratories in Japan for kindly offering the software developed in her laboratories for the purposes of the high-variability training studies. Also, Dr. Sari Ylinen of the Cognitive

Brain Research Unit (CBRU), Institute of Behavioural Sciences, University of Helsinki, for allowing the use of recorded stimuli for the purposes of various research studies of this thesis. I would also like to express my gratitude to the schools that happily collaborated with me, and all participants who volunteered time and effort to be part of this research.

I would like to thank Dr Elizabeth Wonnacott (Oxford University) for her useful suggestions, insightful ideas and comments on earlier drafts of this thesis, as well as inviting me to present part of this research to colleagues in Oxford which was an inspiring experience for me. Many thanks to Jackie and George Ewan for their willingness and patience to proof read the entire thesis. Are you sure you have now grown an interest in this topic?

A special thank you to my dear family and friends for their encouragement and support. I owe gratitude to my husband, Dr Elias Karapanos, for being such a great encourager to me. He supported me every way possible, from being a keen participant, showing me smart ways to analyze data, discussing my ideas for setting up new experiments, even keeping me company during endless study nights. His ceaseless encouragement, patience and support have always been a reminder of what a treasure he is in my life.

I thank my best Friend who is always faithful.

Soli Deo Gloria.

Table of Contents

List of Figures

List of Tables

Chapter 0

Preface

Over the past few decades, second language (L2) learning has become an increasingly important part of the core education a student receives in the modern world and particularly in Europe (e.g. Beacco & Byram, 2003; Byram, Gribkova & Starkey, 2002; Boyd & Rozendal, 2003; Cummings, 1991). For a native speaker of a language, the speech sounds of that language are already perceived and established, even by the end of the first year of life (e.g. Kuhl, 2004, 2008). By the age of five, a child is already a competent speaker of his or her native language (L1) (e.g. Kuhl, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999a,b). During school years, a child develops reading and writing skills in their L1, and (in most European countries) is also instructed in the learning of a L2 as part of the main education curriculum. The learning of a second language (L2), however, can be a challenge for the L2 learner, particularly when one takes into account the relation between the learner's L1 speech sounds and those speech sounds of the target L2. The acoustic similarities and differences of the L1 versus L2 inventories can be an important and contributing factor for successful L2 learning (e.g. Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Best, McRoberts, & Sithole, 1988; MacKay, Meador, & Flege, 2001; MacKay, Flege, Piske, & Schirru, 2001; Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006; Hojen, & Flege, 2006). The exact nature of contributing factors is still being discovered, and there is some evidence (e.g. Ylinen et al., 2010; Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann, & Siebert, 2003; Iverson, & Evans, 2007b) that certain acoustic cues might inhibit or facilitate an L2 learner depending on their L1 background.

This thesis wished to explore the contributing factors in L2 learning further, with a view to also examining the area of plasticity in L2 learning. This thesis consists of 3 main research strands covering the following research questions:

1) The first research strand examines cross language effects in speech perception with particular emphasis on cue-weighting effects. Previous research has suggested that L2 learners pay attention to acoustical cues that are not used by native speakers for correct speech perception (e.g. Holt & Lotto, 2006; Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann, & Siebert, 2003). This thesis sought to investigate these ideas further by using paradigms which tested whether acoustic cues that were not used by native speakers were the basis of L2 learners' perception. Specifically, the research questions posed here are: a) to determine whether L2 learners utilise the same or different acoustic cues for perception of L2 sounds as native speakers of the target language do, and b) to investigate aspects of interference between the L1 and the L2 for L2 speech perception.

2) The second strand also further sought to examine whether any cue weighting effects are malleable in training to a more native-like perception, and if so, whether such training effects are affected by age. This is of tremendous theoretical importance, as several speech models (e.g. Flege, 1981, 1987, 1991a, 1992, 1995a; Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; and others) specify age of acquisition effects in a more 'native-like' perception of non-native phonemes. The research questions posed here are a) to what extent perceptual training alters perception (in terms of cue-weighting) of L2 speech vowel sounds, and b) to determine the degree of plasticity in child versus adult L2 speakers in regards to training.

3) The third strand examines phonological processes in L2 reading. The research question posed here is how L2 orthography could affect L2 phonological processing, especially in relation to spelling differences between the L1 and L2? Previous research has highlighted the relation

and interaction between reading and phonological processing (e.g. Landerl, Wimmer, & Frith, 1997; Pattamadilok, Knierim, Kawabata-Duncan, & Devlin, 2010), and phonological processing skills have been reported to facilitate reading processing (e.g. Adams, 1990; Badian, 1998; Bryant, 1986; de Jong & van der Leij, 2002; Denton, Hasbrouk, Weaver, & Riccio, 2000). Also, Hulme, Hatcher, Nation, Brown, Adams, & Stuart (2002) relate phonemic awareness to word reading ability. These assumptions could suggest some theoretical implications for L2 learners: given that L2 learners are listeners, speakers, as well as readers in both their L1 and L2, any phonological or orthographic interference between the two languages could be affecting L2 learners on a phonological and/or reading level. The underlying orthography of each language is proposed here that may be a contributing factor to speech perception outcomes for the L2 learner (Landerl *et al*., 1997; Ziegler, Stone & Jacobs, 1997). For example, if orthographic cues are used in the L1 in a way that they aid phonological processing in the L1, the L2 learner could erroneously apply the use of similar orthographic cues in an analogous pattern in the L2 which could even affect L2 speech perception outcomes.

This thesis focuses on Greek learners of English for a number of reasons: Firstly, there are an interesting combination of auditory and/or orthographic contrasts in Greek versus English – of particular note is the fact that the Greek language provides an opportunity to study a language where duration is not phonemically relevant and therefore contrasts with other studies that study languages where duration is phonemically relevant (e.g. Ylinen *et al*., 2010, see section 2.1 for a review). Also, differences in transparency levels between Greek and English (see Section 1.6.3 for a more detailed discussion) can aid the investigation on the possibility that differences in orthographic transparency between a L1 and L2 may affect the perception of phonemes in their written form, which is proposed to be a contributing factor for phonological processing outcomes. Thus, examining L1 Greek and L2 English can provide interesting insights on the interaction of phonological and orthographic processes given the underlying differences of the two language systems. Secondly, despite an abundance of studies with other language combinations (e.g. Finnish-English (e.g. Ylinen *et al.*, 2010), Spanish – English (e.g. Rivera-Gaxiola, Garcia-Sierra, Lara-Ayala, Cadena, Jackson-Maldonado, & Kuhl, 2012; Cerviño & Mora, 2009; Escudero, 2000; Fox, Flege, & Murno, 1995), Italian – English (e.g. Flege, 2002; Flege & MacKay, 2004), German – Spanish (e.g. Iverson & Evans, 2007a), Japanese – English (Hattori & Iverson, 2007; Lengeris, & Hazan, 2007; Hazan, Sennema, & Faulkner, 2004; Hazan, Sennema, Iba, Faulkner, 2005) to name a few), this is one of the very first studies to date (alongside with recent research work by Lengeris, 2009; Lengeris & Hazan, 2010) to consider Greek (L1) – English (L2). Thirdly, there is a practical and applied need for this work as English is taught as the main second language in educational institutions throughout Greece.

The ideas for the first strand of this thesis stem from theories of L1 development and learning mechanisms. As infants become selectively tuned to native sounds, language-general abilities gradually become language-specific (e.g. Best & McRoberts, 2003; Werker & Tees, 1984; Werker & Lalonde, 1988; Kuhl, Stevens, Hayashi, Deguchi, Kiritani, Iverson, 2006; Tsushima, Takizawa, Sasaki, Shiraki, Nishi, Kohno, Menyuk, Bast, 1994). So, learning a L2 might be difficult or near impossible. Because L1 sounds are already established, this raises the strong possibility of L1 – L2 interaction or even L1 transfer (e.g. Bohn, 1995; Strange, 1998). A number of L2 learning models have been proposed (e.g. Perceptual Assimilation Model (PAM), Best 1994, 1995a, 1995b; the Speech Learning Model (SLM), Flege, 1981, 1987, 1991a, 1992, 1995a; the Native Language Magnet model (NLM), Kuhl, 1991; Kuhl *et al.*, 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995); the Perceptual Interference model (PI), Iverson, *et al.*, 2003) that describe (to varying degrees) possible mechanisms by which new phonetic categories are acquired in L2 speakers. It has been suggested (e.g. Iverson, *et al.*, 2003) that the acoustic cues that L2 learners use in practice are often different to those used by L1 speakers, thereby leading to inferior performance on perceptual categorization tasks. In other words, L2 speakers have been reported to attend to non-critical or secondary cues rather than critical cues as the native speakers of that language do. This suggests a mechanism by which L2 learners rely on acoustic cues that are perceptually identifiable to them but not so relevant when it comes to a 'native-like' classification (e.g. Iverson, *et al*., 2003; Ylinen *et al*. 2010). These issues are further discussed in Section 1.2.4.

To investigate this, Experiment 1 (Chapter 2) explores a specific English vowel contrast (i.e. /i/ vs. /ɪ/) perceived by Greek L2 English learners and English native speakers. Manipulation of the auditory stimuli allowed the exploration of cue-weighting effects (a detailed discussion on cue-weighting is provided in section 2.1), to determine the degree to which non-phonemically relevant cues were used by L2 speakers of English. Adult and child learners were tested in order to explore how age and L2 experience could be affecting L2 speech sound perception. These experiments consider L2 learning theories and implications, and discuss how acoustic information of English vowel contrasts are perceived by Greek L2 learners of English and how the L1 (Greek) versus different levels of L2 experience could be affecting speech perception performance. Experiment 3 (Chapter 3), investigated the possibility of whether the cue-weighting effects seen in Experiment 1 were influenced by consonantal cues surrounding the vowel contrast under consideration. This idea stems from previous research that has suggested the possibility of consonants surrounding a vowel having an effect on the vowel quality (de Jong & Zawaydeh, 2002). In order to test this possibility, a similar experimental design was used as in Experiment 1, but in this paradigm any existing consonant cues were ambiguous in an experimentally controlled fashion (a detailed account is provided in Section 3.2.2.1). Experiment 3 did not observe any consonantal cues influencing perception of the target vowels as predicted by de Jong & Zawaydeh (2002)'s earlier work. The outcome, however, of these experiments confirm previous studies that suggest L2 learners' attention to cues or features that are non-critical for speech (vowel) sound discrimination by native speakers of the target language (e.g. Iverson, *et al*., 2003; Bohn, 1995).

The second strand of this thesis explores brain plasticity, particularly examining age of acquisition-effects in relation with L2 experience. The brain is a dynamic mechanism with incredible learning abilities, which appear to be heavily influenced by age (e.g. Kuhl, 2010, 2004). The main research question here is to explore the link between the brain's early learning capacity and L2 experience, following perceptual training intervention. In other words, do child L2 learners with little L2 experience have an advantage over adult L2 learners with much longer L2 experience?

Experiment 4 (Chapter 4) explores these issues using a high-variability perceptual training study (HVPT; see section 1.4.2 for an overview). This perceptual training intervention was aimed at identifying whether specialised training could shift attention to the primary features (or cues) for correct vowel identification. Although some previous studies have attempted variants of cue-weighting with little success (e.g. Morosan & Jamieson, 1989; McCandliss, Fiez, Protopapas, Conway & McClelland, 2002), there has been some positive progress in this regard in more recent years (e.g. Ylinen *et al.*, 2010). It was interesting to observe that all groups showed significant improvement following the training intervention. This was also the first study to train child L2 English learners as well as adults. A dramatic result was shown in that despite the children's initial (pre-training) low scores, their *degree* of improvement outperformed the adults' which would suggest that maturational constraints do exist. Adults also improved, which emphasises the significance of the HVPT programmes, however children demonstrate a more robust capacity for perceptually rearranging their phonetic cue-weighting and therefore appear more flexible in their ability to acquire new phonetic categories.

The third strand of the thesis examines phonological processes via a battery of reading tasks. Experiment 2 (Chapter 3) investigated auditory perceptual identification and discrimination but the orthographic word stimuli presented on the screen were replaced with pictures. It demonstrated 'confusions' in Greek child and adult participants alike: the same participants were tested with pictures and with word stimuli performed worse in the case of pictures, which points to the fact that there may be some orthographic cues embedded in the words which 'aid' perceptual identification. Further study in Experiment 5

(Chapter 5) investigating reading performance by Greek adult participants used a number of reading tasks that included rhyming judgment and homophonic judgment for word and non-word conditions. This purely reading experimental study showed that, indeed, orthographic information included in the vowels of target English words used in these tasks, appear to be valuable 'cues' for phonological processing that results in correct or incorrect reading performance. This suggests that there is a possibility of a parallel in the visual modality to what is found in the auditory modality – that visual or orthographic cues may also be weighted by L2 learners differently to native speakers in the perception of speech sounds of a L2.

Chapter 1

First Language Development and Second Language Learning.

The development of language starts at birth, with infants initially starting to hear language sounds spoken in their environment (Goswami, 2008; Crystal, 1991). By four months of age, babies develop the ability to discriminate speech sounds. They attempt language production through babbling around the age of six months and by nine months infants can produce all basic speech elements (or phonemes) that consist their native language (Hunt & Ellis, 2004). By about four years of age they are able to compose full sentences (Steinberg & Sciarini, 2006).

For infants, this seemingly effortless task is a tremendous feat, especially when one considers all the facets of language that they master (new phonemes, new meanings, intentionality, etc.). What makes language development (as opposed to any other skill) quite unique is the fact that most of the progress (at least for spoken language) is accomplished with little explicit instruction. Indeed, artificial intelligence approaches are still far from producing a language machine that would 'crack the speech code' and learn language following the patterns that the human child brain mechanism is programmed to use (Kuhl, 2004, p.831). Infants are born with the ability to perceive and discriminate phonetic representations of any human language in the world (e.g. Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008; Kuhl, 2004; Maye, Werker & Gerken, 2002; Aslin, Jusczyk, & Pisoni, 1998). However, this ability is gradually shaped through experience into language-specific perception. By exposure to a given language, within the first twelve months they have language-specific abilities for discriminating native to non-native language speech contrasts (e.g. Best & McRoberts, 2003; Werker & Tees, 1984; Werker & Lalonde, 1988; Kuhl, Stevens, Hayashi, Deguchi, Kiritani, Iverson, 2006; Tsushima, Takizawa, Sasaki, Shiraki, Nishi, Kohno, Menyuk, Bast, 1994). Therefore, language development seems to simultaneously follow two divergent paths: the path of language improvement (for native language speech perception) and the path of language decline (for non-native language speech perception) (Kuhl *et al.*, 2008).

Recent research has also highlighted the importance of social interaction for language development even suggesting that language cannot be developed merely based on innate abilities (e.g. Tomasello, 2003; Kuhl, 2004; Yu, Ballard, Aslin, 2005; Yoshida, Pons, Cady, Werker, 2006). Infants are born citizens of a world that is interactive and have an innate need for communication. It is this innate need for communication that serves as an important underlying factor which triggers language development and learning. Therefore, the social environment in which an infant is nurtured has an important impact on language acquisition and development. From a neurobiological perspective, existing social constraints form the way infants learn language as they have the capacity to acquire the regularities of the language input they are exposed to when sociolinguistic interactions take place in their environment (Kuhl, 2004).

1.1 L1 Speech Development

Phonological development involves two aspects: the learning of sounds (and sound combinations) that make up the phonological system of the particular native language an infant is exposed to (i.e. all the speech sounds that are permissible in that language), and the learning of how to produce these sounds in a native-like manner. Every language consists of a set of speech sounds which may be unique properties of a particular language or shared with other languages. The speech sounds that a language consists of are called *phonemes* and are described as the basic primitive speech sounds that make up a language. These can be vowel or consonant sounds. For instance, the difference between the word *bat* and *cat* in English is made by changing only one initial phoneme: */b/* versus */k/*[1].

For normal language development, one of the first jobs of an infant is to be able to discriminate speech sounds from other (non-speech) sounds (e.g. environmental noise such as music or vacuum cleaners). Jusczyk (1997) proposed infants' speech-sound preference over other types of sounds and findings demonstrate that even newborns show preference toward listening to speech compared to non-speech (Vouloumanos & Werker, 2007).

The type of language input that a child receives from birth comprises strings of connected words. Naturally spoken language is far from the utterance of isolated speech sounds. Even *'motherese'*[2] (a form of exaggerated and slower speech used in early interactions with infants) does not provide input of speech sounds in isolation (e.g. Stern, Spieker, Barnett, MacKain, 1983; Kuhl, Andruski, Chistivich, Chistivich, Kozhevnikova, Ryskina *et al.*, 1997; Fernald & Kuhl, 1987; Burnham, Kitamura, & Vollmer-Conna, 2002; Trainor & Desjardins, 2002). Phonological development has been suggested to occur not only as a matter of exposure to language speech input but also due to the social interaction involved (Kuhl, Tsao, & Liu, 2003; Goswami, 2008). Therefore, an infant engages with the task of identifying the boundaries between these basic primitive speech sounds that specific clusters of these speech sounds make up words with associated meaning. Speech sounds that are distinctive to the human brain are an approximation of 600 consonant sounds and 200 vowel sounds (Goswami, 2008). Most languages use a small fraction of these sounds, for example approximately 40 phoneme speech sounds are used in English (Breen, Bowers, & Welsh, 1996; Donovan, 1996; Goswami, 2008).

As the term 'development' suggests, speech development is a longitudinal and ongoing process. During this process the child not only needs to attend to the

---

[1] A more extensive analysis of phonemes, especially in relation to Greek and English, is provided in section 1.5.

[2] Other relevant terms reflecting the kind of speech that infants receive, include '*parentese*' (Steinberg, 1993), '*adult-to-child language (ACL)*' (Reich, 1986), or '*child-directed speech (CDS)*' (Pine, 1994).

speech sounds heard in the surrounding context, but also separate them into individual speech sounds, as well as start producing them individually and gradually combine them into meaningful units (or words) (Keenan, 2002). There are several mechanisms that have been proposed to describe the process of speech perception, and these are detailed in the following section.

1.1.1 Speech Perception

The following sections examine firstly the area of *categorical perception* of phonemes, as it is a fundamental of speech perception. Then, a discussion of how categorical perception can be affected by age is also made.

1.1.2 Breaking down the speech signals: Phoneme perception

Infants start with the ability to discriminate phonetic contrasts of all languages (e.g. Kuhl *et al.*, 2008; Kuhl, 2004; Maye, Werker & Gerken, 2002; Aslin *et al*., 1998). However, through exposure to a (native) language spoken in their environment, it is thought that infants use computational strategies by analysing the 'statistical distributions' of the native language stimuli they hear (Saffran, 2003; Kuhl, 2004). The 'statistical distributions' of speech sounds offer clues for the phonemic structures of the native language (Saffran, 2003; Kuhl, 2004). By 12 months they then start perceiving speech sounds *categorically* as a result of their exposure to their L1 environment and this phenomenon has been described as *categorical speech perception* (Liberman, Harris, Hoffman, & Griffith, 1957; Liberman, Harris, Eimas, Lisker, Bastian, 1961a; Liberman, Harris, Kinney, Lane, 1961b; Hoff-Ginsburd, 1997). The term 'categorical' refers to the fact that for two phonemes (different physical sounds), for instance */b/* and */p/*, there are a number of non-identical sounds that are classified together in one of the two categories. In the case of /b/ and /p/, both are *plosive sounds* (also referred as *voiced* and *voiceless stops*) and the difference between them is the amount (or degree) of vibration of the vocal cords: complete occlusion of the vocal tract is followed by a short cessation of airflow release and for voiceless stops (e.g. /p/) this is perceived as silence whereas for voiced stops (e.g. /b/) this is perceived as brief attenuation of sound (Borden, Harris, & Raphael, 2003). What determines such categorization, in the */b/* versus */p/* example, has been found to be a measurable point at which the similar but not identical */b/* sounds stop being perceived as */b/* and are perceived as */p/*. In other words, rather than being a gradual change in perception from /b/ to /p/, there is a sudden change to perceive the sound in the other category.

Liberman *et al*.'s (1957) seminal work neatly demonstrated this effect. In that study, participants were asked to identify the consonant of 32 instances of synthetic two-formant consonant-vowel (CV) stimuli. Their results showed a remarkable symmetry (see Figure 1.1): stimuli with the most sharply rising transitions were labeled as /b/, stimuli with less rising transitions and with slightly falling frequencies were labeled as /d/, and a third category /g/ was also identified for those stimuli with F2 transitions that would fall more steeply.

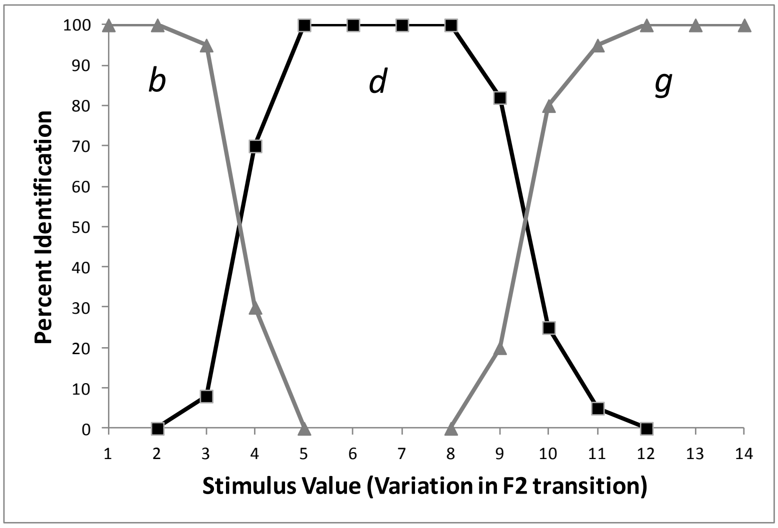


Figure 1.1: Identification data showing the percent of identification when each stimulus was identified as /b/, /d/, or /g/ (Modified from Liberman et al., 1957).

Eimas, Siqueland, Jusczyk, & Vigorito (1971) investigated whether categorical perception occurred in infants (1 and 4 months old) using the sucking method for syllables ‘ba’ and ‘pa’. The ‘ba’ syllable was used as a continuous background sound while the infant was sucking a dummy. There was decline of the rate of sucking when the ‘ba’ sound was played. If the rate of sucking would change at the point when the ‘pa’ sound would be introduced, this would indicate that the infant had perceived a different sound and was responding to it. The stimuli used varied in steps of 20ms in the voice onset time (VOT), with values at: -20, 0, 20, 40, 60, and 80ms. Eimas *et al*. (1971) reported similar results for both infant groups (1 and 4 month olds) with significant changes in sucking rates in the case of stimulus change from +20 ms to +40 ms, while the same absolute magnitude of either -20ms to 0ms or 60 ms to 80 ms VOTs did not initiate any change in sucking rates. Adult results had also showed that the stimulus with a +20ms VOT was perceived as ‘ba’ and the stimulus with a +40ms VOT was perceived as ‘pa’. These results indicate that categorical perception is not only evident for adults but is also demonstrated in 1-month old infants.

In a further study by Werker & Tees (1984), they tested whether categorical perception also occurs in other languages in both infants (6-8 months old, 8-10 months old, and 10-12 months old) and adults (native-Americans and English). Werker & Tees (1984) considered the following contrasts, /ba/ - /da/ (English contrast) and /ki/ - /qi/ (Native-American contrast). Adult participants were instructed to press a relevant key every time they heard a different sound, infants were rewarded every time they turned their head when the sound changed from /ba/ to /da/ or from /ki/ to /qi/. Results showed that both contrasts were perceivable for native-Americans and 6-8 month old English babies, however, for English adults the /ki/ - /qi/ contrast was not perceivable. An interesting pattern for the non-native /ki/ - /qi/ contrast was observed in the case of older English babies: only 57% were able to discriminate at 8-10 months and this dropped to 10% at the age of 10-12 months.

Similar patterns were also observed in the case of Hindi contrasts (Werker & Tees, 1984). Two groups were tested: English and Hindi adult listeners and, English and Hindi-learning infants (aged 6 – 12 months). Two Hindi contrasts were tested: /da/ - /Da/ (Hindi retroflex-dental place distinction) and /t^ha/ - /d^ha/ (Hindi voiceless aspirated and voiced aspirated, voicing distinction). Performance was significantly better for Hindi adults compared to English adults, especially for the /da/ - /Da/ contrast. English infants also discriminated the two contrasts whereas the English adults were not able to discriminate well. Also, following a longitudinal study, native-American babies who appeared to have no difficulty discriminating the /ki/ - /qi/ contrast, between 6 to 12 months of age, showed that they were not so well able to distinguish the non-native contrast. Werker & Tees (1984) concluded that categorization and discrimination abilities for non-native contrasts appear to decline within the first year of life.

It is worth noting that categorical perception is a highly constrained process as infants' high sensitivity appears to be specifically associated with speech sounds as compared to other acoustic sounds that may even share the same physical properties (Kuhl, 2004). It has been argued that the physical changes that occur in phonetic boundaries of different languages are not accidental but it is auditory perceptual abilities that determine these boundaries. It has been suggested that it is mainly the natural auditory discontinuities that influence the phonetic range of the world's languages (Kuhl, 2004, 1991). The process of speech perception for infants, therefore, involves learning of these phonetic boundaries that are associated with their own native language repertoire of sounds. The phoneme boundaries of the particular native language that an infant is exposed to from birth gradually shape the infant's perception to the native boundaries employed (Kuhl, 2004). Thus, language-general perceptual abilities gradually become language-specific (e.g. Polka, Colantonio, & Sundara, 2001; Nittrouer, 2001). This process, occurring around the age of 8-10 months, has been described as a process of functional reorganization rather than entire loss of the sensitivity of language-general patterns (Werker, 1995).

Besides identification and discrimination of language phonemes, it is necessary for infants to also learn to perceptually classify these sounds. Natural speech phonemic stimuli vary according to individual speakers, tone, speech rate, phonetic context and the like. Infants must learn to classify all variations of phonemic stimuli to appropriate categories. The establishment of language-specific phoneme representations for infants leads the way for different-category speech sounds in one language to be classified as same-category speech sounds in another. An example is the /r/ - /l/ contrast that is discriminable for speakers of most languages, for native Japanese speakers the English /r/ is perceptually more similar to English /l/ than Japanese /r/ (e.g. Goto, 1971; Strange & Dittmann, 1984; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, Fugimura, 1975; Takagi & Mann, 1995, Katsura, Flege, Guion, Akahane-Yamada, Yamada, 2004).

Infants, who are exposed to their native language as opposed to a non-native one, tend to gradually specialize in their native sounds and develop patterns that show language-specific effects, they develop high competence at discriminating native phonemes and decreasing abilities to discriminate non-native phonemes. This process points to the existence of some mechanism that allows for such development. Infants appear to develop phoneme categories that serve as 'prototypes' and are language-specific (Kuhl, 1991). The dramatic decline of discrimination between non-native contrasts by one year of age leads to the assumption of the existence of a sensitive period for language acquisition which is discussed in more detail in later section (see section 1.3). These 'prototypes' of native phoneme categories serve as the most representative phoneme types with which other similar-sounding phoneme sounds can be compared.

The learning process that infants follow toward forming these phonetic categories is by establishing distributional frequencies: distributional properties of native sounds that appear in the infant's input are tracked, and frequently occurring acoustic features gradually become properties of the infant's phonetic inventory. Even in the case where the physical distance of different stimuli is the same, non-prototypical sound stimuli are compared to the prototype of this category, causing a 'magnet effect' (e.g. Kuhl, 1991; Kuhl, 1992; Kuhl, 1993a,b; Iverson & Kuhl, 1995). Kuhl, Williams, Lacerda, Stevens, & Lindblom (1992) tested whether this effect occurs only in the case of native prototypes and how native language experience could affect that. They found that American 6-month old infants were successful at discriminating an English vowel prototype /i/ and its variants but showed lower discrimination rates for the Swedish vowel prototype /y/ and its variants; Swedish 6-month old infants similarly appeared to successfully perceive the native Swedish vowel prototype /y/ and its variants but this was not the case with the English vowel prototype /i/. Infants' significantly stronger magnet effect for their respective native language prototype suggests that linguistic experience shapes infants' perception of native speech sounds as early as 6 months of age (Kuhl *et al.*, 1992).

Linguistic experience also appears to be a necessary condition for vowel perception and attention to native language distributional cues (Kuhl, 2004) Infants' sensitivity to distributional cues is such that even limited exposure can initiate learning. Maye *et al.* (2002) tested infants' sensitivity to distributional patterns by exposing two groups of 6 and 8 month old infants initially to a speech sound continuum of voiced unaspirated and voiceless unaspirated consonant stops, and familiarizing them with 8 sounds for approximately 2 minutes. Infants were then exposed to a continuum arranged in two different patterns: a bimodal pattern where frequency distributions were arranged at the ends of the continuum, and a unimodal pattern where frequency distributions were arranged in the centre of the continuum. A preference technique was used (following a familiarization task) and results demonstrate infants' sensitivity to distributional cues supporting the hypothesis of statistical learning patterns at the speech perception process.

Although results from studies mentioned above demonstrate that language learners have powerful statistical learning capacities, it should be noted that the statistical learning mechanism proposed as a pattern for language learning is not necessarily a widely accepted theory and has constraints. For example, research showing that this mechanism is not limited to humans (e.g. Newport & Aslin, 2000) and also the ongoing debate of whether learning is a statistical-based or rule-based process (see Saffran, 2003). Nonetheless, it has been suggestive as an account which would explain part of the infant learning mechanisms proposed by current speech development theories (e.g. Kuhl, 2004; Saffran, Aslin & Newport, 1996a).

1.1.3 Word segmentation

Having broken down ambient language into basic primitive speech sounds, infants are then required to learn how these speech units (or phonemes) are synthesized into words that convey meaning. When considering the synthesis of English phonemes alone, there are over half a million words that can be created. Unlike written language, spoken language – as mentioned already - does not offer the spacing cues that indicate word boundaries (Cole & Jakimik, 1980). Although acoustic pauses are present in natural speech, these pauses do not necessarily mark word boundaries. Therefore, for the infant language learner there should be some mechanism that accounts for this implicit learning. Considering the variability of word structures available in the world's languages, an innate mechanism could not account for this issue alone but a more likely approach is the contribution (and interaction) of both the perceptual and learning mechanisms of the infant learner and the language input available (Saffran, Newport, & Aslin, 1996b).

In a seminal study, Jusczyk & Aslin (1995) tested infants' word segmentation process. A phrase containing a target word was presented in fluent speech to 7.5 month old infants using a head-turning method. A familiarization task was followed by presentation of stimuli of either target words (i.e. words already presented in the speech input) or novel words (i.e. words that did not appear in the speech input). Items were played while the infant maintained the head-turn position toward the source of the sound stimulus. Significant head-turning time differences were observed for the cases of target versus novel word presentations. These results indicate that 7.5 month old infants are able to discriminate between familiar and novel word stimuli. The same task was also performed by 6 month old infants who failed to show different results between the two word categories suggesting that it is possible for this skill to be developed in the period between 6 and 7.5 months of age, or further exposure to the stimuli is required for 6 month old infants in order for results to match the 7.5 month olds' performance.

A similar study (Thiessen & Saffran, 2003) that allowed for more familiarization time with target words for 6.5 – 7 month old infants demonstrated successful results, while 6 month old infants showed high rates of word segmentation only with the addition of cues (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). Also, 9-month old infants show preference in listening to phonetic patterns that are present in their native language while 6-month old infants do not appear to show such preference (Jusczyk, Cutler, Redanz, 1993a; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993b).

But, what is that mechanism that allows infants to segment words that exist in the speech input? Saffran *et al*. (1996b, p.607) propose that it is through the use of 'distributional cues' that infants are able to segment word boundaries. These distributional cues refer to statistical regularities that exist within sequences of syllables. This proposal could lead to word segmentation strategies that are language-independent and according to Thiessen & Saffran (2003) it appears that infants use such cues earlier than other cues (e.g. phonotactics (Mattys, Jusczyk, Luce, & Morgan, 1999), allophonic variation (Jusczyk, Hohne, & Bauman, 1999a), stress patterns (Morgan, Bonamo, & Travis, 1995; Jusczyk, Huston, & Newsome, 1999b), or effects of coarticulation (Johnson & Jusczyk, 2001). Seidl & Johnson (2006) proposed the *Edge Hypothesis* which claims that it is the edges of an utterance that provides infants with the necessary information toward locating words in the speech stream. This is a different proposal to previous claims (e.g. Jusczyk *et al*., 1999b; Mattys & Jusczyk, 2001; Saffran, Aslin, & Newport, 1996a) that it is utterance-medial positions that carry necessary cues for word segmentation. Morgan & Saffran (1995) tested 9-month old infants on rhythmic regularities in a number of basic syllabic arrangements. Results indicate that 9-month old infants used prosodic and distributional cues during speech segmentation processes. A general assumption is that infants appear to make use of a number of cues (e.g. sequential and suprasegmental information) available in the speech stream that enables them to integrate such information toward segmenting word units.

1.1.4 Language development beyond the first year of life

Infant related research, as discussed above, proposes that speech perception abilities decline after the first year of life and also there is a shift from language-general to language-specific perception patterns (e.g. Kuhl, 2004; Werker & Tees, 1984). These assumptions create the impression that little, if any at all, perceptual development occurs beyond the first year of life. However, there are three main reasons to expect perceptual advancement during the subsequent childhood years: 1. the substantial vocabulary growth (Anglin, 1993) which should be considered to initiate changes in the processing of speech patterns; 2. continuous exposure to the native language can have significant influence in the learning of a second language (L2) (e.g. Walley & Flege, 1999); 3. higher level phonological representation tasks (such as reading, writing) that children engage with and master during school years is expected to play an important role on phonological representation and processing (Goswami, 2000). Therefore, language learning continues beyond an infant's first year of life. The learning of a second language has also been the norm for child learners over the past decades in particular. Section 1.2 discusses the abilities and constraints that are realized by second language learners and how the development of their native language could have a specific impact on L2 learning.

1.2 Second Language (L2) Learning

The term second language (L2) learning refers to the process of conscious learning of a second language in addition to the speaker's native language or, according to Flege (1988, p.226, crediting Locke, 1983), L2 speech learning is the process by which the language learner becomes able to 'articulate or perceive a speech sound differently after (as compared to before) massive exposure to a foreign language'. The age at which a second language is learnt classifies the learner as an early (or child) language learner or late (or adult) learner. In the area of second language learning, two main issues have attracted particular attention: firstly, the age at which one starts learning a second language and secondly the extent to which second language learning is affected by the speaker's first (or native) language. This section discusses second language learning reflecting on these two central areas, focusing on spoken second language in particular.

A number of models have been suggested to account for L2 acquisition. This section discusses four models which have been considered the most influential paradigms in current cross-language speech perception research, namely the *Perceptual Assimilation Model* (PAM) (Best 1994, 1995a, 1995b), the *Speech Learning Model* (SLM) (Flege, 1981, 1987, 1991a, 1992, 1995a), the *Native Language Magnet* model (NLM) (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995), the *Perceptual Interference* model (PI) (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann, & Siebert, 2003). A very recent model, the Automatic Selective Perception (ASP) model (Strange, 2010), has also been proposed but has not been extensively empirically tested.

The PAM is primarily concerned with adult listeners' perception of the sounds of a second language, and the development of first language speech perception in infants. The SLM discusses the development of perceptual categories as adult L2 learners become more familiar with L2 sounds and makes predictions on the production of L2 speech sounds. The NLM proposes a *Perceptual Magnet Effect* where vowels are being perceived as closer to their prototypes than might be suggested by acoustic properties alone, suggesting "reduced discrimination sensitivity" (Kuhl & Iverson, 1995, p.131) in the case of two L2 vowels being close to a single L1 vowel, where the listener's ability to discriminate L2 vowels is hindered already by established L1 vowel prototypes. The PI model, an extension of the NLM, suggests that category learning may be hard when listeners are overly sensitive to irrelevant acoustic variation as a result of L1 experience. The recent ASP model (Strange, 2010) predicts the difficulty levels of cross-language speech perception based on existing phonological and phonetic similarities and differences between the L1 and L2. It proposes that native speakers of a language use an automatic selective perceptual mechanism for processing of L1 phonetic contrasts and this is due to L1 experience. In contrast, for L2 learners attentional resources are necessary in order to discriminate these L2 contrasts.

It should be noted here that other L2 perception models have also emerged. For example, the Phonological Interference Model (PIM) (Brown, 1998, 2000) which discusses how L1 influences are generated when L2 segments are acquired and the degree of phonological knowledge involved, the Ontogeny Phylogeny Model (OPM) (Major, 2002) that describes the underlying principles of the L2 phonological pattern formation, and the Second Language Linguistic Perception Model (L2LP) (Escudero, 2005) which claims that the initial state of learners' L2 perception is equivalent with the speech perception mechanism employed by native speakers of the target language (c.f. Best & Tyler, 2007), and postulates that the degree of non-native perception can be a reliable indicator of future L2 development. These models, however, are not discussed further as the emphasis in this thesis is more phonetic rather than phonological. Also, the ASP being a relatively new model, it is still a working model that remains to be extensively empirically tested and therefore this thesis will focus on the four prominent L2 speech perception models (i.e. PAM, SLM, NLM, and PI).

1.2.1 PAM

The Perceptual Assimilation Model (PAM) (Best 1994, 1995a, 1995b) argues that phonetic categories of the learner's native language (L1) are used to handle new language categories and the model's central hypothesis is that listeners *assimilate* non-native sounds to the native sounds they perceive as most similar exemplars. The PAM has been described as a 'fundamentally gestural model' (e.g. Banich & Mack, 2003, p.338) as it considers that perceptual primitives form articulatory gestures. Specifically, perceptual similarity is described in terms of dynamic articulatory information involved, i.e. the way articulatory features form the speech signal. The model predicts that a non-native speech sound that is assimilated to a native category will be a good (but non-ideal) or notably deviant example of that category (Best, 1995a, 1995b). Best (1995a) argues that only a few non-native sounds (e.g. clicks for speakers of non-click languages) are likely not to be assimilated as speech sounds (Best, McRoberts, & Sithole, 1988), while assimilation will occur for almost all non-native speech sounds. Thus, the degree of L1 – L2 similarity determines the degree of assimilation of L2 sounds into L1 sound categories. The model also proposes that non-native sound discrimination accuracy levels relate to the way L2 sounds have been assimilated to L1 sounds. A reason for such lack of efficiency in categorizing non-native speech could be the attunement to language-specific features. Tests where L2 listeners are asked to categorize or give category-goodness ratings to non-native sounds can assess assimilation (Best, 1995a). The PAM proposes a set of assimilation patterns according to which a pair of non-native speech sounds will be discriminated (Best, 1994, 1995a, 1995b), which are described in detail below.

Two-category (TC) assimilation

This pattern occurs when constituents of an L2 contrast are assimilated into two native sound categories. When this occurs, excellent discrimination rates are predicted. A number of studies have tested and confirmed this assumption: Best & Strange (1992) tested Japanese listeners on an English syllable-initial /w/-/j/ continuum and found evidence of assimilation to the Japanese /w/ and /j/, with discrimination being at the same level as native English listeners. Polka (1992) also found assimilation between Salish uvular and velar ejectives (/q'/ versus /k'/) to Farsi voiced uvular and velar plosives (/G/ and /g/) by Farsi listeners who reached native-like perception standards. English listeners were also observed to assimilate the French /y/ versus /oe/ to English /u/ versus /ʊ/, and assimilated French /oe/ and /ə/ to English /ʊ/ and /ʌ/ with high discrimination levels (Best, Faber, & Levitt, In: Best, 1995b). Finally, a study by Polka (1991) also found assimilation between initial position Hindi voiceless-unaspirated dental and retroflex plosives (/t̪/ versus /ʈ/) with English /ð/ versus /d/ by English listeners with high discrimination levels. Moderate discrimination levels were observed for Hindi breathy-voiced dental and retroflex plosives (/d̪ʰ/ versus /ɖʰ/) which were assimilated to English /d/ and /t/.

Single-category (SC) assimilation

In this case, poor discrimination is predicted when two sounds are equally deviant from the native sound, but then are assimilated to one native category. An example is the Japanese listeners' assimilation of an English syllable-initial /r/-/l/ continuum to Japanese /w/, which shows poor identification levels (Best & Strange, 1992). Also, English listeners assimilate the initial position Hindi prevoiced dental (/d̪/) and retroflex plosives (/ɖ/) to the English /d/, and Hindi voiceless-aspirated dental (/t̪ʰ/); retroflex plosives and (/ʈʰ/) were also assimilated to English /t/ with discrimination at chance levels (Polka, 1991). Norwegian high front out-rounded and unrounded vowels also appear to be assimilated to English /i/, as reflected by poor identification levels (Best, Faber, & Levitt, In: Best, 1995b).

Category-goodness (CG) difference

In this scenario, each constituent of an L2 contrast assimilates to a single native category but one constituent L2 contrast is more deviant from the native sound than the other. In other words, they differ in the degree to which they are good examples of that category. Discrimination, in this case, is predicted to be moderate to very good. Evidence comes from a number of studies, for example, Best & Strange (1992) observed Japanese listeners' assimilation for an English syllable-initial /w/-/r/ continuum to the Japanese /w/ but the English /w/ was observed to be a better match and identification levels were good. In a study by Polka (1992), English listeners assimilated Farsi /ɢ/ and /g/ to English /ɡ/. Those who reported that the Farsi /ɡ/ was a better match reached native-like perception rates. Good discrimination levels were observed for English listeners who assimilated German /y/ and /u/ to English /u/ with German /u/ as a good match and German /y/ a poorer match (Polka, 1995).

Both uncategorised (UU)

It is suggested that the existing similarity between the new (L2) and native (L1) sounds determines whether the two sounds will be assimilated as speech sounds but without belonging in the same native sound category. In this case, discrimination levels are varying. In support of this hypothesis, Polka (1992) found assimilation of the Salish /q'/ and /k'/ by English listeners although discrimination was easier on the basis of the following vowel or other acoustic cues, such as duration, suggesting that these sounds had not been categorised. Best & Strange (1992) tested goodness fit and discrimination for the English /r/ and /l/ contrast by Japanese listeners. Both sounds were found to be poor fits for any Japanese sound in the Japanese inventory and discrimination was also poor suggesting that both sounds were uncategorized.

Uncategorized versus categorized (UC)

One sound may be assimilated to a native category and the other may be assimilated as speech sound without being assigned to a native category. Discrimination levels are expected to be good. For example, for the Farsi /G/ and /ɡ/ contrast, English listeners perceived the Farsi /ɡ/ as the English /ɡ/ but the Farsi /G/ was perceived as a low-back vowel with no initial consonant (Polka, 1992). This was suggested to stand as evidence for the UC distinction (Best, 1995b). However, in a study by Guion, Flege, Akahane-Yamada, & Pruitt (2000) showed that although the English /s/ was found to be a good fit for the Japanese /s/, the English /θ/ would fall between two sound categories: the Japanese /s/ and the [ɸ] (i.e. an allophone of /h/). Discrimination in this case was poor suggesting that the PAM may need to be revised in order to account for such results (e.g. an uncategorized sound being in close approximation to a sound that has been categorized).

Non-assimilable

In this case, neither sound is assimilated as speech sound and, depending on the acoustic differences between the relevant sounds, high identification levels are expected. In a study by Best *et al.* (1988), Zulu clicks were not assimilated as speech sounds but instead English listeners described them as 'clicks', 'plops', 'pops', 'finger snap', 'water drips', 'tongue clucking' and other similar descriptions (Best *et al.*, 1988, p. 352) with discrimination rates reaching 80% up to 99% (depending on the acoustic differences involved between exemplars of each pair). They argued that listeners were able to perceive the acoustic differences because the Zulu clicks had not been assimilated to speech sounds that would have 'blocked' listeners' perception. Considering vowel speech sounds, Best (1995b) argues that they cannot be classified as non-assimilable contrasts given the gesture based articulatory space available for the production of speech sounds, especially when considering the fact that certain vowel categories (such as, low vowels, stops, labial and velar articulations) are universal.

1.2.1.1 PAM-L2

L2 speakers often have an observable accent that may also reflect difficulty at perceiving L2 speech sound differences that signal a change in meaning (e.g., "pat" vs. "bat"). Previous studies (e.g. Nobre-Oliveira, 2007) have shown that training can improve perception levels, with long-term benefits, however discrimination levels may not necessarily reach the accuracy levels of a native listener. An extended version of the PAM (Best, 1995a), PAM-L2 (Best & Tyler, 2007), predicts the possibility of perceiving L2 distinctions and how a language-specific phonological system can change with extensive L2 experience. One of the PAM-L2 predictions is that L2 learners can learn to perceive L2 segments but the level of success can vary. The model makes predictions on how adult L2 learners incorporate L2 sounds in their phonetic system and also how language-specific experience (in this case the L1) can influence perception of phonemic contrasts in the L2 (Best, 1994; Best & Strange, 1992). In the case of L2 minimal contrasts being perceived by L2 learners, the following possible outcomes are predicted:

1. The L2 learner can perceive one of the two L2 phonological counterparts as an equivalent (thus, perceptually assimilated) L1 phonological category. In this case, discrimination difficulty will be minimal for the L2 learner.

2. The L2 learner may perceive both phonological L2 contrasts as equivalent to a single L1 phonological category but one category being a better exemplar as compared to the other that may be classified as a deviant sound. Although this deviant would be initially learnt as a variant phone of the assigned L1 category, through experience or training it is thought that a new phonological category can be developed to accommodate the deviant.

3. It is also possible that both L2 phonological contrasts are perceived as either good or poor equivalent sounds of a single L1 category. The L2 learner is predicted to assimilate the two L2 sounds phonetically and phonemically into a single L1 category and discrimination levels will be very poor as the minimal pair (words with only one contrasting sound) is likely to be perceived as a homophone pair (i.e. of identical pronunciation) rather than contrasting.

4. Both L2 sounds can be perceived as speech categories without being assimilated to any L1 categories. In this case, the learner can establish one or two new phonological categories; depending on whether the distinction between them is perceived clearly.

The PAM/PAM-L2 (Best, 1995a; Best & Tyler, 2007) supports the view that perceptual learning can still occur in adulthood, but it emphasises the differences between child and adult L2 learners in their ability to perceive non-native sounds. In this respect, a greater emphasis is placed on external factors (such as the amount of L1 vs. L2 experience for the formation of L2 phonetic systems), than on internal factors (such as maturational constraints due to maturational changes in the brain). In this respect, adults could be considered as disadvantaged in their perceptual learning of L2 speech sounds due to their greater ‘history of experience’ with their L1, compared with child L2 learners

(Best, 1995, p.198). The PAM/PAM-L2 also does not make clear predictions about the formation of two phonetic systems when these are acquired simultaneously as in the case of bilinguals. The PAM/PAM-L2 has largely been applied to cases where exposure to the L2 occurs after the L1 has already been established (e.g. Guion *et al*., 2000).

1.2.2 SLM

The Speech Learning Model (SLM) developed by Flege (1981, 1987, 1988, 1991a, 1992, 1995a) deals primarily with the phonological acquisition of L2 segments by adult speakers. It examines both perception and production, testing the effect of experience of non-native speech sounds. It claims that inaccurate perception of L2 sounds is a main reason of foreign accented speech. It proposes that phonetic similarity between non-native and native sounds is what determines their classification as ‘similar’ or ‘new’ depending on L2 sound equivalence with native sounds (Flege, 1987). The production of an L2 sound will eventually correspond to its perception-based phonetic category, but inaccurate perception of the L2 sounds may lead the learner to establish non-accurate phonetic categories of the L2 which would serve as the base for L2 production. The more dissimilar to the native phoneme categories the L2 sounds are, the more likely it is for new phonological categories to be developed and produced accurately. It also emphasizes the amount of L2 experience, considering the L2 learner’s *Age of Arrival* (AOA) or *Length of Residence* (LOR) effects which make it more of a longitudinal approach to L2 learning.

Little or no correlation was observed between length of residence (LOR) in the L2 context and the degree of foreign accented speech in the case of L2 adult learners (Flege, 1988). A comparison between residence in the L2 context for 6.8 and 0.8 years did not show significant results (Flege, 1991b). However, individual variation seemed to be a significant factor for different groups with LOR differences. For vowel production, such individual differences may appear and even stabilize within the first six months of exposure to the L2.

Thus, the properties of the L2 categories will be established and will demonstrate major changes within the first months of L2 learning. However, it has been proposed (Flege & Liu, 2001) that the LOR does aid predictions for L2 perception in the case of interaction with native speakers of the L2 and significant amount of exposure to the L2. For instance, perception of English word-final consonants was measured for Chinese listeners. It was found that perception correlated with LOR for participants whose L2 learning involved extensive interaction with English native speakers compared to those with relatively little interaction with English native speakers.

In terms of categorization of L2 sounds, Flege (1987, 1991a, 1992, 1995a) used the term ‘equivalence classification’ for ‘similar’ or ‘identical’ to the L1 sounds. Although the acoustic differences may be perceivable by the L2 learners for what is classified as ‘similar’ sounds, the equivalence classification process prevents them from making use of them in speech production and speech processing in general. L2 sounds that have been classified as similar are assimilated into a diaphone, i.e. a sound category to account for both the L1 and L2 sound (Flege, 1995a). Also, the SLM claims that the L1 phonetic categories established in early years continue to develop throughout a person’s life.

Continuous exposure to the L1 and L2 can modify the phonetic properties of the developed diaphone category. This would require the detection of within-category discrepancies plus the ability to assign speech sounds to different phonetic categories. Flege & Hammond (1982) found evidence for perception of within-category differences: monolingual speakers were able to perceive and produce a number of within-category differences but were assigning foreign accented speech sounds to relevant phonetic categories. This indicates that L2 phonetic categories of native-speaker norms cannot develop due to the equivalence classification effect. Therefore, production of L2 speech sounds will involve intermediary properties of L1 and L2 ‘similar’ sounds. The equivalence classification effect may thus limit the L2 learner’s approximation to the L2 speech sounds.

An example is a study by Munro (1993) that found tendency to produce English vowels with spectral properties that reflect intermediate sounds between native-English and ‘similar’ native Arabic vowel sounds by Arabic learners of English. In terms of perception, Bohn & Flege (1990) found that the English /i/ and /ɪ/ contrast was perceptually similar to the German /i/ and /ɪ/ contrast while Bohn & Flege (1992) confirmed this by testing production of the English /i/ and /ɪ/ contrast by German speaking English learners. Analysis of the spectral and duration values revealed properties of an intermediary segment compared to the respective native German and native English speech sounds. The same results were observed for the English /ɛ/ sound and its ‘similar’ German /ɛ/ counterpart. Properties of an intermediate rather than native English or native German sound were detected even for experienced learners. Therefore, equivalence classification effects seem to block the emergence of sound categories of L2 norms.

The SLM also proposes that it is possible to develop ‘new’ sounds which are not subject to the equivalence classification effect. It is possible for L2 learners to develop new sound categories but these categories may be based on different or differently weighted cues compared to those used by native speakers (e.g. Flege, 1995a). Also, in the case of bilinguals, the SLM hypothesizes that they maintain phonetic contrasts between the entire inventory of sounds for both the L1 and L2 as they have single phonetic space for both languages. Thus, what determines the properties of a new sound category, is the establishment of phonetic category dissimilation so as these categories could be deflected.

When testing the French /u/, which is classified as ‘similar’ sound, experienced native English learners of French produced second formant (F2) values that were intermediate between the native English and native French sound, while production of the ‘new’ French /y/ sound showed F2 values that matched production by monolingual French speakers (Flege, 1987). Flege, Schirru, & MacKay (2003) examined English vowel production by Italian immigrants to Canada with early and late *age of arrival* (AOA) rates. Production samples of the English /e/ by participants with late AOA showed less formant movement than when produced by English native speakers, showing that the English /e/ was perceived as a similar sound to the Italian /e/ which differs only by the lack of formant movement. Participants with early AOA (who made little use of their native language) produced the English /e/ with formant movement that was significantly greater than that produced by English native speakers. This indicates that the English /e/ sound had been perceived as a new category and the exaggerated formant movement was an attempt to maintain the phonetic contrast between the Italian /e/ category and the new English /e/ category.

The SLM model includes the term 'identical' (as distinct from the terms 'similar' and 'new') and states that 'a "U" shaped rather than a linear function may describe better the effect of varying differences between L1 and L2 sounds' (Flege, 1992, p. 566). It also claims that it is possible for a L2 sound to be pronounced accurately if it is identical to a L1 sound, or the similarity between an L1 and L2 sound can be so close that any existing differences can go unnoticed in case the L2 sound substitutes the L1 sound. A main problem of the SLM model is determining what comprises sounds as 'new' or 'similar' as it is hard to distinguish the boundary between the two categories and there is no principled method to determine that distinction. The SLM is based on the degree of perceived phonetic dissimilarity (as formulated in Flege, 1995a) and proposes that the greater the perceived dissimilarity, the greater the possibility of a new category to be developed. Flege (1991a) suggested that in an attempt to create an accurate picture of what comprises sounds as 'new' or 'similar' requires phonetic distance scaling methods by participants in a laboratory setting. Although Flege (1995b) considered the different approaches on how the phonetic distance should be measured (e.g. in terms of acoustic difference, the phonological features, or articulatory gestures involved), no definitive conclusion was reached.

Flege (1992) stated that an L2 vowel would be perceived as new if it appeared within the vowel space that was 'uncommitted' in the L1 vowel space. Rochet (1995) claimed that those L2 sounds that are assimilated are perceived as similar to the L1 sounds and therefore this leaves no 'uncommitted' perceptual space; it is the perceptual boundaries that convey more important information than perceptual prototypes as the SLM suggests, and as vowel categories expand so as to reach the boundaries of the next category this leaves no room for uncommitted space or the establishment of new vowels. For instance, in the case of Catalan English and Brazilian Portuguese, listeners did not appear to have uncommitted space for the Parisian French /y/ although they were able to discriminate a high vowel continuum and classify them into an /i/ or /u/ category (Rochet, 1995). However, according to the SLM (as proposed in Flege, 1995a), L2 sounds do not fall within uncommitted (assimilated) perceptual space in the case of a 'new' category being established. In terms of the underlying learning mechanisms, the SLM does not account whether they are general or specific or whether perception is due to the extraction of linguistic (i.e. phonetic) or acoustic information.

The SLM includes a quite extensive set of assumption and hypotheses, however the main features of the model could be summarized in the following: 1. the same processes of L1 acquisition are also used in L2 acquisition regardless of the L2 learner's age; 2. phonetic categories established early in life as L1 speech sounds continue to develop throughout life; 3. greater perceptual dissimilarity between L1 – L2 speech sounds can result in the formation of a 'new' speech category, while perceptual similarity between L1 - L2 speech sounds will result in a single category which will link the two (L1 vs. L2) speech sounds and this will also be resembled in production; 4. perceptually identifiable differences between L1 and L2 speech sounds are less likely to be identified by the L2 learner as age increases.

1.2.3 NLM and Neural commitment

The *Perceptual Magnet Effect* and *Native Language Magnet Model* (NLM) (Kuhl, 1991; Kuhl *et al.*, 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995), offers a complementary account for why listeners may be unable to perceive spectral variation within the vowel space. It emphasizes that experience shapes the perception of the physical world. More specifically, although there may be very clear acoustical differences between two sounds, a non-native speaker will not hear them unlike a native speaker of that language. The NLM proposes that future learning is constrained by initial mappings and neural commitment, i.e. the L1 language-specific filter constrains the acquisition of L2 speech sounds. Although this language-specific constraint does not relate to a critical period, it is argued that with increasing age, neural structures become more 'committed' and the resulting mental mappings are claimed to interfere with the creation of new mappings based on L2 input.

The speech perception process is assumed to involve general mechanisms for auditory perception that process the speech input as acoustic rather than phonetic (or linguistic) information. The NLM suggests that at the early developmental stages listeners establish acoustic prototypes for their native phonemic categories. These prototypes have magnet-like effects (as the name of the model also suggests), i.e. the perceptual space that is in phonetic approximation to each of these native prototypes is attracted by the established category, posing difficulty in the later perception and discrimination of phonetic variation of prototype categories as compared to non-prototypes. Unlike the SLM, the NLM predicts asymmetrical discrimination levels and lack of acoustic experience causes a failure to establish new prototypical categories.

A study showing asymmetrical discrimination levels was conducted by Grieser & Kuhl (1989) where infants and adult native speakers performed well at identifying prototypical sounds of native speech categories. English native speakers were asked to rate the prototypicality levels for more than 100 /i/ tokens and also choose a most representative prototypical and non-prototypical exemplar (see also Kuhl, 1991). Six-month-old infants and adult native listeners were tested in their perceptual ability to distinguish exemplars from variants, where for the variants the F1 and F2 were altered, equidistant in *ms* and were centered on each exemplar. The two groups were tested in terms of their ability to discriminate the exemplars from the variants. Variants with the same distance from their exemplars were equated with the prototypical (vowel) sound compared with the non-prototypical for both groups. The Perceptual Magnet Effect attempts to explain this by suggesting that linguistic experience causes distortion to the perceptual space surrounding the prototype category and that results in reduced discrimination sensitivity.

According to Kuhl *et al*. (1992), exposure to particular languages creates perceptual magnets: for example, a magnet effect was observed for six-month old Swedish infants who were tested on the Swedish /y/ but no magnet effect was observed for the English /i/; interestingly, no magnet effect was observed for American infants on the Swedish /y/ but for the English /i/. In a study by Nakai (1997, 1999), Japanese and Greek listeners were asked to listen to a grid of synthetic vowels and select those that sounded as closest variants to their native vowels, i.e. /i/, /e/, /a/, /o/, /u/, as possible. In all cases, with /a/ being the only exception, selected vowels were those cases with extreme spectral properties than the average F1 and F2 values compared with their native vowel productions. It is suggested that the Perceptual Magnet Effect may be due to vowel extremity and not due to the fact that a native category has been established.

The Perceptual Magnet Effect is closely related to the idea of the existence of native prototypes. However, the NLM also suggests and emphasizes the existence of boundaries. Kuhl & Iverson (1995) suggest that boundaries are considered to be innate properties. They argue that infants learn how to ignore boundaries that are innately specified and are not relevant to their L1. Those areas that have been restricted by the remaining boundaries are then dominated by the magnets. Boundaries do not disappear but are still accessible in the case of non-assimilated sounds, such as non-speech. The NLM proposal that perceptual L1 magnets are established attempts to explain why L2 vowels that are in close spectral approximation to L1 vowels are hard for L2 listeners to discriminate. This is in agreement with the PAM/PAM-L2 assumption in terms of assimilation of a L2 vowel to L1 vowels and also the SLM predictions of equivalence classification.

An expanded version of the NLM model (i.e. NLM-e) was proposed by Kuhl *et al.* (2008) suggesting that neural commitment affects plasticity levels and determines to what extent the phonetics of a second language are to be learnt later in life. Figure 1.2 graphically presents an overview of the NLM-e (Kuhl *et al.*, 2008).

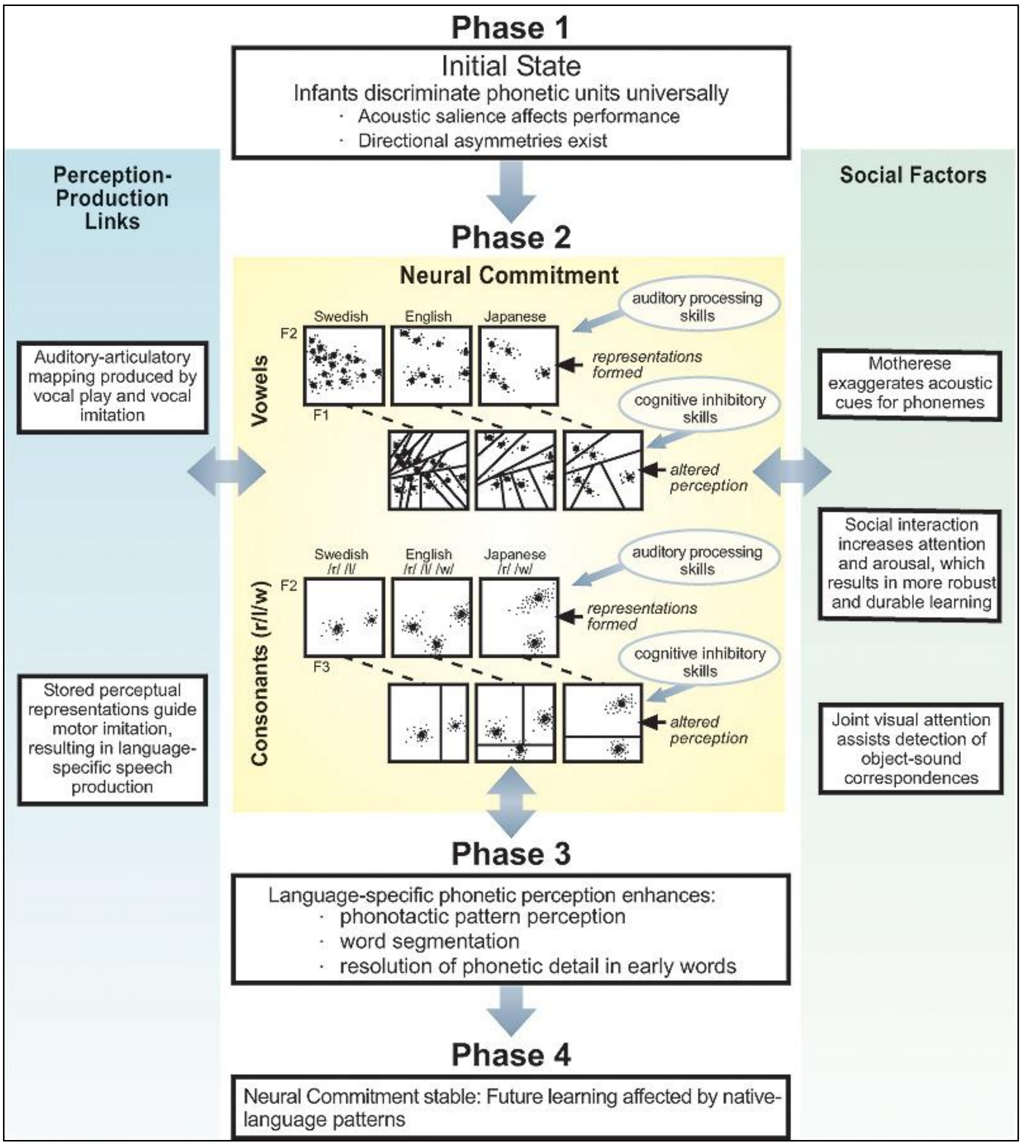


Figure 1.2: Native Language Magnet theory expanded (NLM-e) (adopted from Kuhl *et al.*, 2008). Phase two includes data from studies on Swedish (Fant, 1973), English (Dalston, 1975; Flege *et al.*, 1995; Hillenbrand *et al.*, 1995) and Japanese (Iverson *et al.*, 2003; Lotto *et al.*, 2004).

Although infants are born with the ability to potentially learn any world language (Phase 1, Initial State), age poses a restriction on this ability. As it is predicted by the NLM-e, infants' sensitivity to phonetic cues declines and social interaction seems to play a significant role, leading the infant to understand language within its surrounding social environment (Tomasello, 2003). Other factors include exaggeration of speech sounds (especially vowel sounds, so called 'motherese' speech) and distributional properties of speech sounds that interact and most likely lead to infant perceptual change and neural commitment processes taking place (Phase 2). Although this is described as a 'phase' in the model's terminology, it is not easy to predict the exact stage when such perceptual shift takes place as this is highly dependent on the language environmental input that is available. The search for phonetic cues available in the environmental language input leads to enhancement of the ability to detect language-specific phonetic cues while the ability to detect or search for language-general (or non-native) sounds appears to diminish (Phase 3). This is the stage when language-specific attunement is proposed to take place (Kuhl *et al.*, 2008). Although early phases of the NLM-e propose that during infancy neural networks are not completely formed and therefore allow for unrestricted learning, later (i.e. adulthood, Phase 4) exposure to a novel language does not necessarily form neural structures as a result of L2 exposure. Neural commitment appears more stable and L2 learning is affected by L1 patterns.

A general assumption of the NLM-e deals with the 'degree of plasticity' that may exist at the stage of L2 phonetic learning as existing perceptual representations (and the degree of neural commitment) can affect L2 learning (Kuhl *et al*., 2008, p.992). This assumption postulates why adult L2 learners have greater difficulty to acquire and discriminate certain L2 speech sounds: perceptual similarity of the L2 speech sound to an already established (prototypical) L1 sound can determine whether the L2 sound will be 'pulled' toward the L1 sound category.

However, the perceptual magnet effect proposed by the NLM-e cannot satisfactorily account for language learners with similar dual-language experience (e.g. bilinguals) and thus similar amounts of exposure to 'prototypical' or 'non-prototypical' speech sounds (e.g. Fowler, Sramko, Ostry, Rowland, & Hallé, 2008; Mack, 1989). Also, it does not account for the amount of change on the 'degree of plasticity' (Kuhl *et al*., 2008, p.992) given, for example, laboratory based perceptual training (see section 1.4 for an overview) or generally increased L2 exposure (e.g. L2 learner immigrating to a country where the L2 is spoken, an aspect that is accounted for by the SLM with Age of Arrival (AoA) rates). Therefore, it might be necessary to also identify the existence of learner-specific prototypes rather than language-specific ones. Such task, however, would require to also identify learner-specific variables which determine how or when a prototype is formed and established which could be a rather impractical solution. These aspects of the NLM-e demonstrate some methodological implications in terms of applying this model to L2 learning contexts, but it nevertheless gives a plausible explanation on why some assimilation effects (as proposed by the PAM/PAM-L2 discussed earlier) are often 'attracted' by native categories rather than non-native.

1.2.4 PI (Perceptual Interference)

The Perceptual Interference (PI) account, proposed by Iverson *et al*. (2003), discusses how it is possible for early experience of language to hinder later learning of L2 phonemes. In particular, it points to the fact that acoustic cues that are used by L2 learners may be different to those cues used by L1 speakers and therefore performance can be distorted in perceptual categorization tasks. This results in L2 learners identifying and relying on non-critical cues and consequently failing to identify cues that are critical for the native speakers of the target language. The initial language-general perceptual and cognitive abilities shift to become language-specific causing high levels of difficulty to adults attempting to learn a second language. The PI account is rather complimentary to previously discussed models as it offers an explanation in terms of what is acoustically perceived by the native versus non-native ear.

Supporting data point to the fact that non-native speakers appear to be more acoustically sensitive to cues that are not relevant to the native speaker (Iverson *et al*., 2003). This, however, is suggested to depend on established language-specific categorizations as these may lead to different acoustic cue sensitivity that is closer to that of the native speaker. Therefore, this language-specific dependency for L2 perceptual processing could be a factor that predicts L2 speakers' performance and the degree of sensitivity to critical or non-critical acoustic cues.

An example is the case of Japanese adults tested on the English /r/ - /l/ categorization. Japanese speakers appear to be sensitive to second formant (F2) values of the /r/ - /l/ contrast compared with English native speakers who do not base their perceptual categorization on this particular (F2) acoustic cue but rather on the F3 formant. In an experiment conducted by Iverson *et al*. (2003), Japanese, American, and German monolingual speakers were tested. Identical stimuli of /ra/ and /la/ were varied only in terms of F2 and F3 formant frequency at the point of consonant closure, i.e. F2 and F3 formants were altered generating a two-dimensional scale of equally distributed frequencies following the *mel* scale (Stevens, Volkmann, & Newman, 1937). Identification and goodness tests required the participants to identify each stimulus exemplar with regard to a native phoneme as well as mark it within a 1-7 scale (where 1 was a 'bad' and 7 was a 'good' exemplar of a relevant native phoneme). Results were mapped onto a Multi-Dimensional Scale (MSD) (Kruskal, 1964) showing perceptual space. Also, those examples that were reported as perceptually similar were placed at one side and perceptually dissimilar stimuli were placed at the opposite end of a two-dimensional scale. Results showed that language experience is a main factor for identifying those perceptual spaces.

This becomes apparent when considering cue weighting in the case of different monolingual speakers: although American participants showed sensitivity to the F3 formant values in the case of categorizing the /r/ - /l/ contrast, Japanese participants appeared more sensitive to the F2 formant changes which is not a relevant cue for the native speaker (in this case American participants) (Iverson *et al*., 2003). Interestingly, German participants' performance was more similar to the American than the Japanese participants', suggesting that the equivalent /r/ - /l/ contrast in German more closely resembles the English contrast than anything within Japanese (which does not have an equivalent contrast at all, but instead its own flap[3] consonant) (Iverson *et al*., 2003). This suggests that the underlying native language can determine sensitivity toward L2 cues that may or may not be relevant to the native speaker. Also, the fact that Japanese adults appear 'mistuned' (Iverson *et al.*, 2003) to hearing the English /r/ - /l/ contrast suggests that there is perceptual interference for the boundaries of those perceptual contrasts, therefore blocking Japanese listeners from forming correct category representations and misleading them toward relying on acoustic cues that are perceptually identified but not relevant when it comes to classification.

The PI account complements the theories discussed earlier in this chapter that perception of speech sound categories interfere with earlier experiences leading to a developmental decline in the ability for language learning (Flege, 1995; Kuhl, 1994, 1998, 2000; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002). Exposure to a particular language in relation to another (e.g. the case of Japanese versus German as discussed above) could determine the degree of perceptual interference of non-native contrasts and therefore loss of sensitivity to certain non-native contrasts (due to the language-general to language-specific transition during infancy) may be irreversible in later years when maturation factors are involved. Listeners, however, may still be able to become 'attuned' to L2 contrasts but as a result of 'self-reinforcement' (Iverson *et al*., 2003) that leads to sensitivity to acoustic cues that are not necessarily the ones that would be potentially used if such attunement had occurred earlier in life through relevant language exposure. Although maturation factors do pose barriers toward the formation of new sound categories for adults, they seem to still allow the development of some

[3] A 'flap' consonant is produced with a single muscle contraction in a way that an articulator (e.g. the tongue) is thrown against another; it is similar to a brief 'stop' however it does not employ any buildup of air pressure at the place of articulation and therefore there is no burst release as in the case of a 'stop' consonant (Ladefoged & Maddieson, 1996).

categorization pattern, even if not relevant to that employed by a native speaker.

1.2.5 General Discussion on L2 learning

The four models of speech perception described in this chapter portray a number of similarities as well as a number of distinctive features. Table 1.1 summarizes the four models.

Table 1.1: Summary of the four perceptual models (partially adapted from Escudero, 2005).

	PAM/PAM-L2 (Best, 1995a; Best & Tyler, 2007)	**SLM** (Flege, 1995a)	**NLM/NLM-e** (Kuhl, 1991, 1992, 1993b, Kuhl et al., 2008)	**IP** (Iverson et al., 2003)
Initial learning stage	L1 categories	L1 categories	L1 neural mappings	L1 categories/cues
Learning mechanisms	Same as in L1	Same as in L1	Same as in L1	Interference with L1 cues
Perceptual Development	Reorganisation of categories	Creation of new L2 and/or mapping on same L1 categories	New L2 categories	Sensitivity to critical or non-critical acoustic cues
Prediction	Speech category assimilation	Speech category formation/ merging	Creation of new L2 categories	Acoustic cues of L1 to interfere with L2
Final learning stage	Depending on L1 vs. L2 articulatory differences	Depending on Age of Learning & L2 experience	Depending on L1 vs. L2 experience	Depending on degree of L1vs.L2 acoustic interference

All four models approach L2 speech perception as a language-specific process as it is based on the already learnt L1 speech categories. Also, all models suggest that it is through the learner's interaction with the L2 that speech categories are created or perceptually reorganized. More specifically, Best's PAM/PAM-L2 (Best, 1995a; Best & Tyler, 2007) suggests that it is possible for L2 learners to perceive a non-native contrast with exposure to L2 input where there can be reorganization of assimilated category perception in the L2 (Best & Strange, 1995). However, the PAM/PAM-L2 model does not account for for a specific learning mechanism that may allow such reorganization of assimilation patterns to perceptually take place, however it may be assumed that L1 learning mechanisms account for category learning.

The NLM/NLM-e model suggests that L2 learners can create new speech categories of L2 sounds (although this can be affected by already established L1 mappings of language patterns). However it does not clarify what mechanism may be involved for such creation of new categories. Kuhl (2000) suggests that for L2 categories to be created later in life for the L2 learner, a different mechanism to that used for L1 acquisition may be required, but no specific detail of such type of learning mechanism relevant to L2 development is proposed. Therefore, it is unclear how it is possible for these new categories to be established later in life, especially with possible L1 interference. Nevertheless, it is suggested that although unrestricted learning may diminish after infancy, some later neural commitment may still be possible –depending on the phonological relationship between L1 and L2 sounds which may also determine the 'degree of plasticity' (Kuhl *et al.*, 2008:992) that may be realized at that learning stage.

The SLM deals almost exclusively with the final stage of learning, therefore it does not elaborate on earlier stages of L2 speech perception. This model assumes that L2 adult learners have access to the same learning capacities used by infants or children for L1 acquisition in the sense that the already established L1 categories develop over the life span. It is thus possible for adults to accurately perceive L2 speech sounds as well as perceptually establish new categories (e.g. Flege & MacKay, 2004). It is suggested that the degree of cross-language phonetic differences determine the degree of development or discrimination of L1 versus L2 sounds by L2 adult learners, and this may determine the learner's development in the L2. Therefore, new L2 categories can be formed depending on the perceived phonetic similarity or dissimilarity between the L2 sound and the perceptually closest L1 sound. The more perceptual dissimilarity the more likely it is for new sound categories to be formed (e.g. Flege, 2003). For similarly-sounding sounds between the L1 and L2, the model suggests that these can be treated as similar or same category, which points to the PAM/PAM-L2's assumptions of perceptual assimilation.

The PI is more complementary to previous proposals and could be treated as an extension to the NLM-e model as it concentrates on a particular aspect of L2 perception, the sensitivity to critical (or relevant) or non-critical (or irrelevant) 'cues' (or features). It points to the fact that speech sounds carry a number of cues, but some are more critical than others in terms of sound identification or discrimination. The cues however that are treated as critical by L1 speakers may be ignored by L2 learners who may attend to other non-critical cues in speech identification and discrimination tasks. It suggests that this creates interference which can determine the direction of perceptual cue-weighting for L1 speakers and L2 learners alike.

Although the models described have a number of similarities, they approach L2 speech perception from slightly different perspectives and emphases. L2 speech perception is a complicated process, and so it could be highly optimistic to create a single model that aims at an exhaustive solution, especially taking into account the multiple cross-language combinations that can exist. So it is suggested that these models are complementary and all contribute to account for the various perceptual and neural mechanisms of learning a language, although still not providing a complete account for L2 language learning processes. In the following chapters, a number of studies are presented taking into account the various L2 speech learning approaches, and aim to shed more light on this highly complex cognitive process, i.e. L2 speech perception.

1.3 Sensitive (or Critical) Periods for Language Acquisition

The above sections have provided an overview of language acquisition and speech perception in normal conditions: when an infant is born healthy and grows up in an environment where a language (i.e. a native language) is spoken and therefore is exposed to linguistic input. The learning of a second language has also been discussed in detail with a number of models that have emerged in relation to L2 learning (e.g. the PAM/PAM-L2 by Best 1994, 1995a, 1995b, Best & Tyler, 2007; the SLM by Flege, 1981, 1987, 1991a, 1992, 1995a; the NLM /NLM-e by Kuhl, 1991; Kuhl *et al.*, 1992, 2008; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995; the PI by Iverson *et al.*, 2003; and, the ASP, by Strange, 2010).

The language development stages described in earlier sections discuss the gradual progress that takes place while the infant 'tunes in' the native language, so from a language-general state progresses into a language-specific state (e.g. Kuhl, 2004; Kuhl *et al.*, 2006; Kuhl & Iverson, 1995; Kuhl *et al.*, 2003; Saffran, 2003). Such language development assumes typical social experiences, linguistic input and stimulation. However, a number of cases where infants were deprived of typical social experiences and stimulation, thus not consistently exposed to a significant amount of language (or no language input at all) in the early years of their life, led to the proposal for the existence of 'Critical or Sensitive Periods'.

Penfield & Roberts (1959) first recognized the existence of a critical period in language learning (also, Birdsong, 1999; Herschensohn, 2007; Singleton, 2001). Lenneberg (1967) was the first to introduce the term 'critical period hypothesis' or CPH. This proposal assumes that it is necessary for language to be acquired between infancy and puberty, if it is to be acquired at all, thus this time frame is considered as 'critical'. The critical period assumption was based on the idea that at the end of puberty there are significant maturational changes in the brain that prevent the acquisition of high level skills such as the learning of a language. Later research has proposed the alternative term 'sensitive period', which views the time period between infancy and puberty as a ‘sensitive’ rather than ‘critical’ period, suggesting that language is most easily acquired during this period. However, after this period, it is still possible for language to be learnt but it is considered more difficult and ultimately less successful. Knudsen (1999, 2004) made a clear distinction between a ‘critical’ and a ‘sensitive’ period for language acquisition arguing that neuronal connections are susceptible to environmental input during the ‘sensitive’ period, but later experience continues to influence neural development, unlike the stricter assumptions of a ‘critical’ period. Knudsen (1999) suggested that a critical period may be exhibited particularly in the development of speech perception and early exposure to a language can have indelible effects (e.g. Flege, Yeni-Komshian, & Liu, 1999).

Although evidence for the existence of a sensitive period for language acquisition comes from cases of deprived or feral children who failed to develop language due to lack of early exposure to linguistic input (e.g. Curtiss, 1980; Curtiss, 1977; Goldin-Meadow, 1982), or cases of deaf children’s and adults’ acquisition of American Sign Language (ASL) (e.g. Woodward, 1973; Mayberry, Fischer, & Hatfield, 1983; Curtiss, 1988; Newport, 1984; Newport & Supalla, 1992), this evidence cannot satisfy a proposal for absolute timelines for language acquisition. Not only in some cases deprived children managed to learn language to some degree (e.g. Curtiss, 1980; Curtiss, 1977; Lane, 1976), but further evidence comes from neuropsychology where it is suggested that adults, well beyond the sensitive period, are more likely to suffer permanent language impairment from brain injury compared with children. This is suggested to be due to early-life capabilities of neural reorganization (e.g. Birdsong, 2006; Weber-Fox & Neville, 1999) and brain lateralization (e.g. Lenneberg, 1967; Long, 1990), although Hertz-Pannier, Chiron, Jambeque, Renaux-Kieffer, *et al.* (2002) have challenged the classical age limits for language acquisition imposed by the critical period assumptions. The main argument of the critical or sensitive period is the importance of linguistic input at an early age, preferably from the early stages of infancy. Therefore successful language learning is heavily dependent on early language exposure (e.g. Elman, 1993; Kuhl, 2004).

The proposal of the existence of a sensitive period for language acquisition also extends to the learning of a second language (L2). The main argument being that older learners of a second language rarely achieve native-like fluency compared with fluency often achieved by younger L2 learners (e.g. Chiswick & Miller, 2008; Scovel, 1988, 2000, 2006; Long, 1990; Johnson & Newport, 1989). A wide range of proposals have been suggested based on the existence of a critical or sensitive period for L2 acquisition: L2 learners cannot reach a native-like stage (e.g. Scovel, 1988), other suggestions include age-specific boundaries beyond which native-like fluency is not possible to be achieved (e.g. Long, 1990; Johnson & Newport, 1989), to the extreme line of language learning mechanism deteriorating from birth therefore L2 proficiency is suggested to be unattainable (Hyltenstam & Abrahamsson, 2003). A general assumption is that early language acquisition is better but not to the exclusion of successful adult L2 learning (e.g. Chiswick & Miller, 2008). A number of previous studies (although each testing different aspects of L2 learning) have shown that native-like proficiency of L2 learners is possible even past puberty, therefore past the sensitive period for language acquisition (e.g. Urponen, 2004; Marinova-Todd, 2003; Flege, Murray, & MacKay, 1995; Mayberry, 1993; White & Genesee, 1996; Birdsong, 1992; Cranshaw, 1997).

Recent neuroimaging techniques provide evidence for maturational constraints in relation to L2 learning with particular emphasis on the amount of exposure being a significant factor (e.g. Abutalebi, Kappa, & Perani, 2001; Grosjean, Ping, Munte, & Rodriguez-Fornells, 2003) also highlighting that early learners have considerable advantage over late learners for a second language (e.g. Perani, Abutalebi, Paulesu, Brambati, Scifo, Kappa, Fazio, 2003; Neville, Coffey-Corina, Lawson, Fischer, Emmorey, & Bellugi, 1997; Neville, Mills, & Lawson, 1992). The area of phonetic learning has been suggested to be particularly sensitive to age of L2 acquisition (e.g. Bongaerts, Planken, & Schils, 1995), with age of arrival (thus, start of L2 learning) correlating with foreign accent in a study by Flege *et al.* (1999). Proficient L2 learners, however, could be accounted as evidence that L2 attainment after puberty is not restricted. However, if such language attainment cannot be explained by a critical period account then there should be other influencing and contributing factors.

Although it is not the aim of this thesis to extensively discuss critical or sensitive period proposals, it is important to acknowledge the existing literature on this topic as well as recognize the existence of maturational constraints for language learning, especially L2 learning. However, the main question that this thesis addresses, in view of the sensitive period, is to what extent is it possible to train the child or adult brain in order to perceive and ultimately learn specific L2 contrasts? Also, what may be the differences between child and adult learners between an initial and end state within a very specific and measurable time frame of L2 learning? Also, how plastic is the brain for child and adult populations alike?

1.4 Training the Brain: Training Approaches

Considering language development and L2 learning theories discussed above in relation to maturational constraints, especially constraints in relation to phonetic learning and the concept of sensitive periods, it is possible to over-emphasize this concept as a definitive and ultimate closure of a 'window of opportunity' (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1996) rather than language *learning* itself. Native language neural commitment as proposed by the NLM/NLM-e model (Kuhl, 1991; Kuhl *et al.*, 1992, 2008; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995) points to the fact that the learning of a native language results in the formation of neural networks that are dedicated to re-coding the patterns of the native language (e.g. native language speech). This process takes place early in development when the brain gradually dedicates neural networks to the perception of natural spoken language. However, this early coding and neural commitment interferes with new language input in the case of learning an additional language (e.g. L2). In this framework, early learning of a native language could either promote or restrict later L2 learning depending on the extent of L2 language patterns conforming to the already established neural patterns (e.g. Kuhl *et al.*, 2005; Kuhl & Iverson, 1995). It is likely that L2 learners use native-like strategies for processing L2 speech which may constrain accurate identification and categorization of L2 speech sounds (e.g. Zhang, Kuhl, Imada, Kotani, & Kohkura, 2005; Iverson *et al*., 2003).

A number of studies have used training techniques which specifically aim to improve performance and neural efficiency with reported results showing L2 learners' significant improvement but not necessarily to the level of native speakers' performance (e.g. McClelland, Fiez, & McCandliss, 2002; Pisoni, Lively, Logan, 1994; Zhang *et al*., 2005). The following sections give a more detailed account of the most prominent training techniques to date supporting the notion of perceptual training as a useful tool for L2 learning.

1.4.1 Early Techniques

Early training techniques of speech perception (e.g. Pisoni, Aslin, Perey, & Hennessy, 1982; Strange & Dittmann, 1984; Yamada, 1991) have demonstrated the beneficial effects on L2 learning. Although an early training study by Carney, Widin, & Viemeister (1977) aimed at testing whether it is possible for adult monolingual speakers to be trained in perceiving L1 category differences, this led to the use of such a promising technique and its application to perceptual training for L2 learners. Strange & Dittmann (1984) trained adult Japanese speakers who had lived in the US between 5 and 30 months. Pre- and post-training tests used natural minimal pairs for the /r/ - /l/ contrast (e.g. 'lock' – 'rock') arranged in an identification task (i.e. participants hear one auditory stimulus in each trial and their task is to choose between two visual options) and also a synthesized /r/ - /l/ continuum was tested in an identification and a discrimination task (i.e. participants hear two auditory stimuli and their task is to decide whether they sound the same or different). Training included 14-18 sessions that used the synthesized /r/ - /l/ continuum in a discrimination task. Immediate feedback was given following every trial throughout the training sessions. The post-training tests showed improvement in the tasks that involved the synthesized /r/ - /l/ continuum but improvement was not as evident for the tasks that involved the natural stimuli pairs. Two main issues were identified as contributing to this lack of perceptual improvement with natural stimuli: the task-type that was used throughout the training sessions (i.e. discrimination training task) and the low-variability of the stimuli used (i.e. single talker and single context) in the training intervention. It has been suggested that discrimination training tasks could lead attention to non-critical within-category rather than critical between-category differences (e.g. Jamieson & Morosan, 1986). In other words, attention to the right dimension (e.g. Iverson & Kuhl, 1995, 1996; see also, Uther, Singh, Zipitria, & Uther, 2005) is important in order to achieve successful perceptual training on L2 speech categories.

1.4.2 HVPT Studies

Following early training techniques, a new training approach emerged using high-variability perceptual training (HVPT) (Logan, Lively, & Pisoni, 1991; Lively, Logan, & Pisoni, 1993; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Hazan, Sennema, & Faulkner, 2005; Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005; Iverson & Evans, 2007a; Nishi & Kewley-Port, 2007a,b, 2008; Iverson & Evans, 2009; Ylinen, Uther, Latvala, Vepsäläinen, Iverson, Akahane-Yamada, & Näätänen, 2010). This approach highlights the importance of training with natural tokens arranged as contrasting minimal pairs and within a multiple-talker context which is a closer resemblance to natural real-world speech. This technique has been known for successful learning of L2 categories.

Logan *et al*. (1991) trained Japanese speakers in the English /r/ - /l/ contrast. The Japanese participants were studying in the US and had lived in the US for no more than 3 years at the time of testing. The training consisted of 15 sessions, involved identification task-type, and more minimal pairs were used compared with earlier studies, i.e. 68 minimal pairs of the /r/ - /l/ contrast.

Participants were also provided with immediate feedback. The pre- and post-training tests used the same minimal pair stimuli as used by Strange & Dittmann (1984) but these differed from the stimuli used in the training sessions. Two additional tests of generalization were also included in the post-training test. The first test of generalization involved 96 novel minimal pairs of the /r/ - /l/ contrast but this time produced by a new talker (not previously used in any training or test sessions). The second test of generalization used 98 novel minimal pairs of the /r/ - /l/ contrast which were produced by one of the talkers used in the training sessions. Identification tasks were used during the pre- and post-training sessions but no feedback was given. Pre- and post-training test comparisons showed improvement in the minimal pair identification task as well as both the generalization tests highlighting the importance of stimulus variability and the importance of task-type used for the training purposes.

Following Pruitt (1993)'s critique and suggestions of Logan *et al*. (1991)'s study about the need to test talker and phonetic environment variability effects in a training context, Lively *et al*. (1993) conducted a new training study incorporating this proposal in two experimental conditions. In the first, they trained Japanese native speakers the /r/ - /l/ contrast with minimal pair stimuli produced by 5 native English talkers and a variation of three different positions (i.e. initial singleton positions, intervocalic positions, initial consonant clusters). In the second, participants were trained in a larger array of phonetic environments but all tokens included in the training sessions were produced by one English native speaker. Results from pre- and post-training test comparisons revealed that there was improvement in perception of the /r/ - /l/ contrast. However, results from their first experiment (which showed that learning extended to novel words pronounced by a new talker), led to the conclusion that talker variability is an important aspect for successful cross-language training.

In a subsequent study, Lively *et al*. (1994) examined long-term retention of learning the trained /r/ - /l/ contrast by Japanese participants by using the same experimental methodologies but administering a post-test on a later date. This showed that learning outcomes following the training intervention were retained for at least 6 months following training. The high-variability training approach was since used to train further cross-language contrasts. For example, Flege (1995b) trained Chinese native speakers on the English /t/ - /d/, Pruitt (1995) and Pruitt, Jenkins, & Strange (2006) trained native speakers of English and Japanese on the Hindi dental and retroflex stops, also English native speakers were trained on the Japanese vowel length contrasts (e.g. Hirata, 2004; Hirata, Whitehurst, & Cullings, 2007; Tajima, Kato, Rothwell, Akahane-Yamada, & Munhall, 2008), or later studies trained English native speakers to perceive the Mandarin lexical tones (Wang, Spence, Jongman, & Sereno, 1999; Wang, Jongman, & Sereno, 2003).

In a study by Bradlow *et al.* (1997), Japanese speakers were trained on the English /r/ - /l/ distinction in a high variability training programme with feedback. The study looked at the effects of training on both perception and production. Two groups of adult Japanese speakers participated, one group was trained, the other group was not trained but served as controls. Results on perception improvement confirmed earlier findings on the effectiveness of the HVPT technique. Production results were assessed by a group of English native speakers and findings showed that improvement on the English /r/ - /l/ perception had also transferred to the Japanese speakers' production of the trained distinction. A subsequent study by Bradlow *et al*. (1999) not only replicated previous results but also showed long-term retention of the trained distinction after a 3-month period following the training in both perception and production.

Recent studies have also examined vowel contrasts in a variety of cross-language contexts. Nishi & Kewley-Port (2007a) trained 17 adult Japanese native speakers, who were living in the US for no more than 1 year at the time of testing. They were trained on two sets of American English vowels, a set of 9 and a set of 3 target vowels respectively. Participants were arranged into three experimental groups. Six participants (Group 1) received training on all 9 target vowels, six participants (Group 2) received training on the 3 target vowels, and five participants (Group 3) received no training but served as control group. For Groups 1 and 2 that received training, the standard procedure was followed of pre-training test, training sessions, post-training test, and a follow-up retention test after a period of 3 months. The pre- and post-training tests involved 36 real words (arranged as CVC using a variety of consonant contexts) used for generalization tests, and also nonwords (arranged as CVCə using six different consonant contexts). The training stimuli used were the same nonwords used in the pre- and post-training tests. Nine training sessions were administered. The training used identification task-type and provided immediate feedback. In the case of an incorrect answer on any given trial, participants were also given the opportunity to replay either of the stimuli (correct or incorrect) up to 10 times prior to proceeding to the next trial.

Results overall showed improvement in the L2 contrasts across participants of both groups; generalization tests showed that this improvement was also evident in the case of new words and new talkers, and was also retained after a period of 3 months. One main observation, however, of this study comparing improvement of the two groups was that Group 2 (that received training on the 3 target vowels as compared with the full 9 vowel training set used for Group 1), did not show improvement on non-trained vowel contrasts. This highlights the effectiveness of a wider variety of L2 contrasts being trained and suggests that this could allow participants to familiarize themselves with a broader range of dimensions (e.g. duration or spectral) or allophonic variations involved (Nishi & Kewley-Port, 2007a).

In a later study, Nishi & Kewley-Port (2008) trained adult Korean native speakers on American English vowel perception. Participants were residing in the US at the time of testing but for no more than one year. Three training protocols were used: the first was the same as previously used in Nishi & Kewley-Port (2007a), the second used a combination of the 3 and 9 vowels in a way that participants were trained with the 9 vowel set for the first 6 days and then the 3 vowel set for the last 3 days, and a third where the 9 vs. 3 vowel sets used were reversed (6 day training with the 3 vowel sets and 3 day training with the 9 vowel set). The aim was to identify whether the different training protocols would result in different perceptual learning outcomes. Results showed that Korean learners' perception of American English vowels was improved across all three different training protocols. However, the two new protocols introduced in this study did not show any significant additional benefits compared with the training protocol used in Nishi & Kewley-Port's both 2007a and 2008 studies.

A slightly different training protocol was introduced by Iverson & Evans (2009). They trained native speakers of Spanish and native speakers of German on 14 Southern British English vowels. They used only natural tokens of minimal pair contrasts. The training procedure included an adaptive system which would allow half of the trials to be presented adaptively based on the participant's incorrect responses. Following 5 training sessions, results revealed that all participants improved their perception of English vowels and this improvement was retained in 3 month follow-up tests. More specifically the German group showed a greater degree of improvement compared with the Spanish group. Following an additional 10 training sessions for the Spanish group, they reached the same levels as the German group previously had achieved with only 5 training sessions. This particular outcome certainly highlights the role of the L1 for the effectiveness of HVPT techniques.

Perceptual training intervention was also used by Ylinen *et al*. (2010) training 20 adult Finnish native speakers on the British English tense-lax /i/ vs. /ɪ/ vowel contrast (e.g. *heat* vs. *hit)*. Forty-five minimal pairs were used as training stimuli and 19 of these pairs were used for the pre- and post-training tests. Training sessions and pre- and post-training tests used both natural and modified duration stimuli (for more detail on modified duration stimuli, see Section 2.2.1.2). The training sessions included recorded tokens of 4 different English native speakers (2 male and 2 female) and a fifth English native speaker was used for the pre- and post-training tests. Results showed significant identification accuracy for both natural and modified duration word-types used, which emphasizes the beneficial effects on perceptual training through the use of a specialized perceptual training technique.

A recent HVPT study that considered Greek learners of L2 English was conducted by Lengeris & Hazan (2010). They tested both perception and production of L2 English vowels. Particular emphasis was given on observations on individual differences in performance. L2 learners' frequency discrimination acuity and vowel processing ability was tested. Pre and post-training tests and five training sessions were administered. Eighteen Greek L2 English learners participated in all training sessions and ten Greek L2 English learners served as control group and participated only in the pre and post-training tests. Results showed that perceptual training improved participants' perceptual identification of L2 vowels and this also extended to production of these vowels. L2 vowel perceptual identification was tested in quiet and noise conditions and training seemed to improve performance for both conditions, although perceptual identification of L2 vowels showed higher improvement in quiet (20%) than in noise (15%). Although at a group level, there is obvious improvement as a result of HVPT, the strategies for training could be conceivably improved by looking at individual differences and areas that would benefit each individual.

Finally, recent additions include audiovisual perceptual training technique introduced in a study by Hazan, Sennema, Iba, & Faulkner (2005) as well as videogaming paradigm explored in a very recent study by Lim & Holt (2011). Hazan *et al*. (2005) trained Japanese learners of L2 English on the English /v/-/b/-/p/ and /l/-/r/ contrasts. The study emphasized the effectiveness of audiovisual training compared with audio training alone, and how audiovisual training can enhance sensitivity to visual cues available in the case of non-native phonemic contrasts as well as its effects on improved pronunciation outcomes. Lim & Holt (2011) used a videogame paradigm for perceptual training of L2 categories. They also trained the English /r/-/l/ contrast to Japanese learners of L2 English. Using this videogaming technique it was possible to emphasize associations among sound categories, as well as provide visual cues. Participants essentially were the players of a videogame and their task was to respond to videogame characters, based on the rules of the game. There was no overt categorization task or feedback provided, compared with previous perceptual training techniques. A shift toward a native-like cue-weighting was observed. Participants received 2.5hr training per day for 5 days in total, and reached same levels of improvement as in earlier studies where participants had followed explicit perceptual categorization training lasting 2-4 weeks. These recent techniques highlight the importance of visual cues in addition to auditory cues which could be suggested as being jointly effective for successful perceptual training.

The various high variability perceptual training studies overviewed in this section show the importance of such training technique for the L2 learner. It is evident that perceptual training can result in improved auditory perception of L2 sounds. This has remarkable practical implications for computer-assisted teaching and learning. A multiple talker training paradigm seems an effective way of atuning L2 learner's perception to acoustic information that is critical for correct identification of L2 speech sounds. Taking into consideration first and second language development processes discussed in earlier sections and the advantages of the early learner's brain over an adult learner's brain (from a developmental perspective), it would be necessary to also consider applying such promising perceptual training method to early L2 learners (children). Such application could also allow further investigation and findings in terms of developmental abilities and maturational constraints in the L2 language learning domain. Also, it could allow for further examination on the interaction between L1 versus L2 experience and age effects. Further development of such methodologies could also result in considerable applications to language schools and language learning institutions for the benefit of the L2 learner.

1.5 Greek and English phonemes

Given that a central focus of the research in this thesis is to look at cross-language differences in vowel perception between L1 Greek and L1 English speakers (see also Chapter 0), it would be prudent to compare and contrast the vowel inventories of the two languages. Therefore, the next section is devoted to describing and comparing the vowel systems of both languages.

Every language has a distinctive inventory of phonemes. Phonemes are defined as the ‘distinctive speech sounds of a language’ (Strange & Jenkins, 1978, p.126). Within the Second Language Acquisition (SLA) literature, phonemes are also defined as “phonetic categories, the smallest segments of spoken language that combine and contrast to make up the words of the lexicon” (Strange, 1995, p.5).

A number of features categorize the individual phonemes and set them apart from all other sounds of that language. Therefore, phonemes could be said that are contrastive elements. An example is the formation of minimal pairs, i.e. a pair of words differing by one sound, for instance, ‘*pat*’ versus ‘*bat*’, are phonetically different by one phoneme, the initial consonant (/p/[4] versus /b/), however their meaning is distinct. In the case of variants of a phoneme, for example instances of an aspirated /p/ (as in *pit* ([phɪt][5]) versus non-aspirated /p/ (as in *spit* ([spɪt]), are called ‘allophones’ of a phoneme (e.g. O’Grady & Dobrovolsky, 1987; Rogers, 1991). Examples of minimal pair vowel phonemes in English are for example ‘*feet – fit’* where the vowel phonemes used are /iː/ and /ɪ/ respectively[6]. As the empirical work in this thesis focuses on vowels, the discussion is focused on this area rather than consonants.

Vowels are produced by passing air through the vocal tract with no closure or narrowing of the speech organs sufficient to produce audible friction (Hall, 2003; Borden *et al*., 2003). The oral cavity acts as a resonance chamber and modifications to the shape or size of this area create different vowels. This area is described as the vowel space, or acoustic space (Ladefoged, 2001, 2005).

The main acoustic characteristics of vowels involve: duration, the F_0 (voice fundamental frequency), amplitude and formant structure (Fourakis, Botinis, & Katsaiti, 1999). Other factors may also influence the acoustic characteristics of vowels such as the phonemic category, prosodic or suprasegmental factors,

[4] Phonemes are written between forward slashes, as shown in this example, and are represented symbolically by the International Phonetic Alphabet (IPA). IPA characters are used to describe phonemes and allophones throughout this thesis.
[5] Allophones are written between square brackets, in accordance with the literature.
[6] A detailed review of the English vowel phonemes is provided in section 1.5.2.

stress, emphasis or tempo (i.e. speaking rate), as well as discourse factors (Fourakis, 1991; Fourakis *et al*., 1999; Moon & Lindblom, 1994; Sluijter, 1995). Spatio-temporal variability is a common feature in the case of speech production (Nikolaidis, 2003). Phenomena that are often classified as categorical phonological changes, such as assimilation, deletion, substitution and more, could be considered as gradient and this is suggested to be highly attributed to style and speech rate (Nolan, 1992; Holst & Nolan, 1995).

Every vowel sound is characterized by a set of formants that is referred to as its formant structure. Formant structure is a fundamental feature that defines a phonemic category of vowels due to the frequency components involved. Information about the frequency content of each vowel, such as formant movement or duration, is perceptually required in order for humans to distinguish vowel sounds (e.g. Iverson, Smith, & Evans, 2006; Assmann & Katz, 2005; Hillenbrand & Nearey, 1999). Therefore, each vowel has a set of formant frequencies which enable the listener to distinguish one vowel from another, even if two vowels may spectrally be quite close to each other. Each of these formants has specific frequencies, which are measured in Hertz (Hz). The formant with the lowest frequency is called F_1, the second F_2, and the third F_3. Vowels have a number of formants ($F_{1,}$ $F_{2,}$ $F_{3,}$ F_4), however, usually the first two (F_1 and F_2) are reported as most important in determining vowel quality and in terms of carrying enough information or cues for perceptually distinguishable vowels (c.f. Chiba & Kajiyama, 1941; Potter & Steinberg, 1950; Ainsworth & Millar, 1972; Carlson, Granström, & Klatt, 1979).

Formant structure is also related with the articulatory position (or spectral location) of the vowel (i.e. front versus back and high versus low features). F_1 is associated with the high-low phonetic dimensions and F_2 is associated with the front-back phonetic dimensions of vowels. Therefore, low vowels (such as [a]) are expected to have higher F1 values, high vowels (such as [i] or [u]) are expected to have lower F1 values, whereas front vowels (such as [i]) are expected to have higher F2, and back vowels (such as [u]) are expected to have lower F2 values.

The vowel inventory of a language (i.e. the number of different vowel sounds involved) evidently influences the vowel positions in the acoustic space (Fourakis *et al.*, 1999), as well as how different vowels (V) 'combine' with consonants (C) in a way that is as 'economical' as possible when it comes to movements of the tongue at syllable articulation (e.g. VCV or CVC contexts). This phenomenon has been ascribed to 'base-of-articulation' properties considered as language-specific (Bradlow, 1995, p.1916; also Fourakis *et al.*, 1999).

Every language does not possess totally unique vowels but some vowels are common or 'shared' between languages. For example, the /a/ vowel is very commonly used by several different languages such as English, German, Spanish, Italian, Greek, to name but a few. Cross-language comparisons, however, reveal that even for those vowels that are considered as 'shared' between different languages, there are differences in the formant values (F_1 and F_2 frequencies) even though they may be minimal and perceptually perceived as more of an allophonic variation. Thus, vowels may be considered 'similar' but not necessarily 'identical' or 'shared'. In a study by Bradlow (1995), four vowels [i, e, o, u], considered 'shared' between American English, Spanish and Greek, were compared see Figure 1.3).

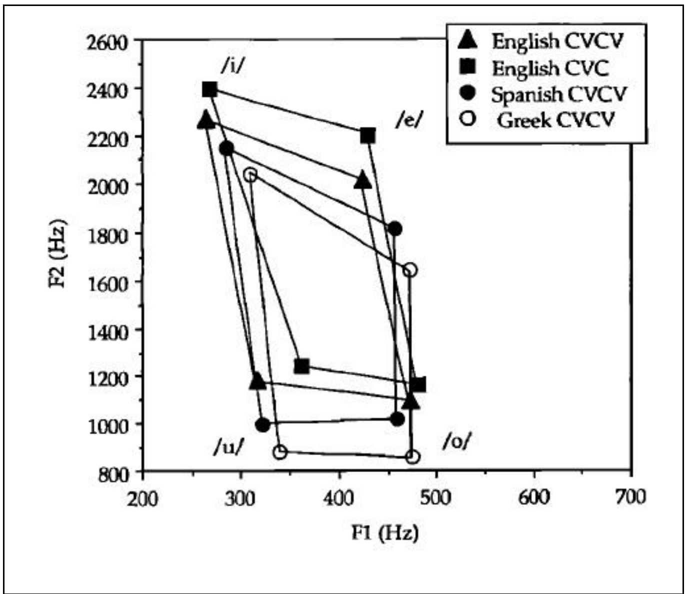


Figure 1.3: Comparison of the F1 x F2 space covered by English (CVC and CVCV contexts), Spanish and Greek (/i/, /e/, /o/, /u/) (adopted from Bradlow, 1995).

Results showed that the F_2 values of the Spanish vowels were lower than the English F_2 values and the Greek F_2 values were lower than the Spanish ones. This concurs with the previously mentioned argument that this could be the result of language-specific 'base-of-articulation' properties (Bradlow, 1995, p.1916; Fourakis *et al*., 1999) so that certain vowels in different languages could occupy similar acoustic spaces but not necessarily share the 'precise phonetic realizations' (Bradlow, 1995, p.1922).

1.5.1 Greek Vowels

This section considers the Standard Modern Greek[7] vowel sounds that underlie the assumptions and predictions made throughout this thesis. The Greek vowel system consists of five monophthongs or distinct vowels [i, ɛ, a, ɔ, u] (Koutsoudas & Koutsoudas, 1962) or [i, e, a, o, u] (Mackridge, 1985; Joseph & Philippaki-Warburton, 1987). This study engages with the description and symbols as suggested by Mackridge (1985) and Joseph & Philippaki-Warburton (1987) for consistency reasons throughout the thesis and due to the fact that they are more widely used in the Greek phonetics literature. Vowels here are considered and discussed particularly based on two features: duration features (vowel temporal length) and spectral features (place of articulation within the F_1 x F_2 space).

In terms of the place of articulation of Greek vowels, a chart depicting the vowel positions for the production of Greek vowels is shown in Figure 1.4. This chart, so called the *cardinal vowel system*, represents a cross-section of the oral cavity: the top portion shows the highest position of the tongue, the bottom portion shows the lowest position of the tongue, the left side represents the front of the mouth and the right side represents the back of the mouth. The

[7] Greek vowels throughout this thesis always refer to Standard Modern Greek or else non-accented Athenian Modern Greek.

cardinal vowel system shown below (Figure 1.4) presents the Greek vowels in their corresponding IPA[8] symbols in terms of their place of articulation.

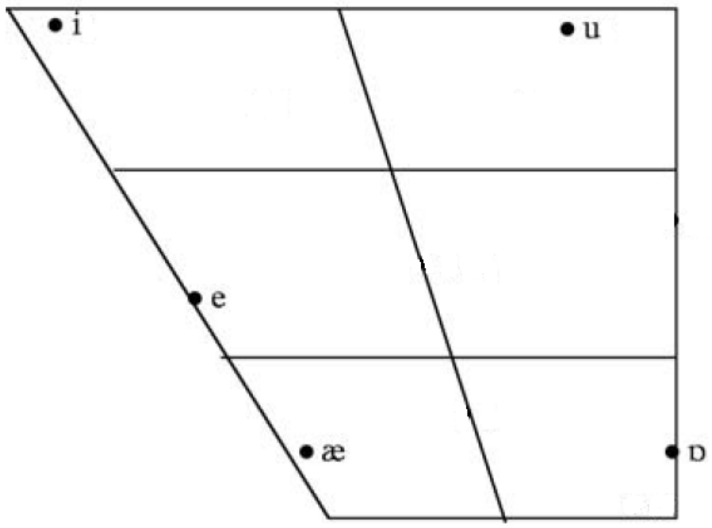


Figure 1.4: Vowel monophthongs of the Greek vowel inventory.

Vowel height refers to the vertical position of the tongue when the vowel sound is articulated and it ranges from 'high' (or close) to 'low' (or open) which also correlates with the first formant frequency (F_1) of the vowel. The horizontal positions of the tongue within the vowel space indicate the 'front' versus 'back' position of the vowel and this is correlated with the second formant frequency (F_2).

Table 1.2: Height and horizontal articulatory positions of Greek vowels.

	Front	Central	Back
High (or Close)	i		u
Mid	e		o
Low (or Open)		a	

[8] IPA (International Phonetic Alphabet).

Table 1.2 presents the Greek vowels in terms of height (frontness and backness) as well as their horizontal position (high - low or close – mid – low) and Figure 1.5a graphically shows the spectral characteristics of the Greek vowels which reveals a correlation with their respective place of articulation.

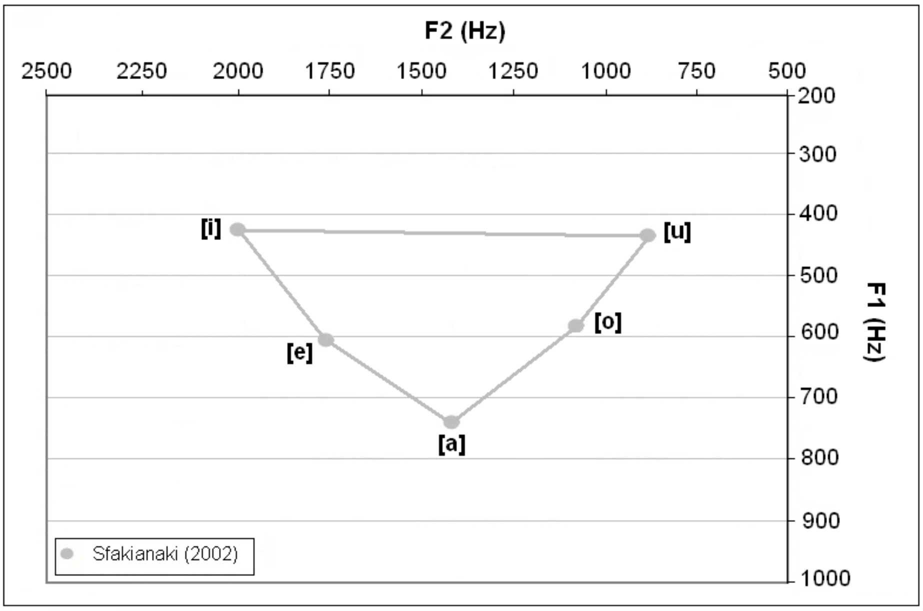


Figure 1.5a: Graphic representation of mean F1 and F2 values for Greek vowels (mean values adopted from Sfakianaki, 2002).

Studies by Fourakis *et al*. (1999), and Nicolaidis (2003) have also provided data on the five Greek vowels along with Sfakianaki (2002), and it is important to point out the formant (F1, F2) variation reported by each of these studies. Figure 1.5b, adopted from Arvaniti (2007), depicts this variation via plotting the data from the three studies on Greek vowels. Although it is not clear why such differences would be present, a plausible explanation is that the experimental design of Nicolaidis (2003) used spontaneous speech data rather than words (or sentences) in isolation. This could be a reason for a rather centralized vowel space reported by Nicolaidis (2003) as opposed to Fourakis *et al*. (1999) or Sfakianaki (2002).

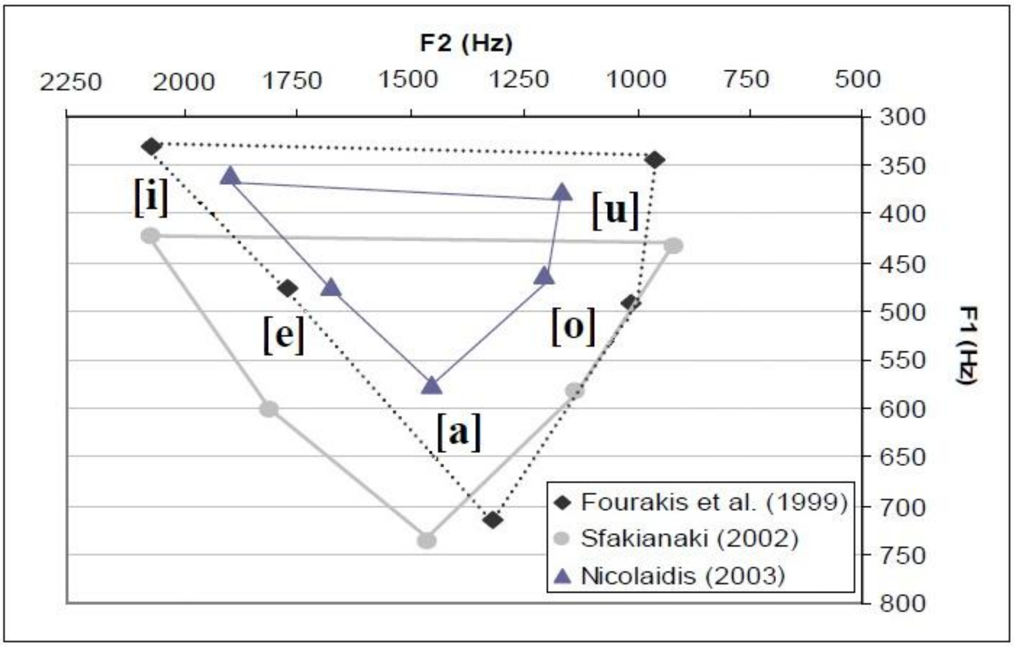


Figure 1.5b: Graphic representation of mean F1 and F2 values for Greek vowels of three different studies on Greek vowels (graph adopted from Arvaniti, 2007).

Perceptually /i/ is a high front vowel and /u/ is high back vowel (Samaras, 1974; Botinis, 1981; Tseva, 1989; Fourakis, *et al.*, 1999; Sfakianaki, 2002; Nicolaidis, 2003; Nicolaidis & Rispoli, 2005; Baltazani, 2007). The two mid vowels /e/ and /o/ are front and back respectively. The low vowel /a/ is intermediate between low and low-mid articulatory position (Nicolaidis, 1991), therefore another appropriate symbol for its phonetic transcription is [ɐ]. Based on the fact that it represents a low central vowel, the /a/ symbol will be used throughout this thesis for convenience in comparing with its English counterparts as well as being consistent with the literature (Mackridge, 1985; Joseph & Philippaki-Warburton, 1987; Koutsoudas & Koutsoudas, 1962). Formant values (F_1, F_2) of all Greek vowels pronounced by Greek adult speakers (age 20-28 years) from Sfakianaki (2002) are summarized in Table 1.3.

Table 1.3: Mean formant values F_1, F_2, F_3 (in Hz) of the Greek vowels produced by Greek adult speakers (values adopted from Sfakianaki, 2002).

	F_1	F_2	F_3
i	423	2073	2593
e	601	1811	2560
a	736	1466	2459
u	434	921	2460
o	583	1137	2479

A study by Jongman, Fourakis, & Sereno (1989) examined stressed vowels produced by 4 male Greek speakers and reported that the Greek vowels are very well distinguished within the acoustic space, therefore contrast between the vowel categories is maximized. Kontosopoulos, Ksiromeritis, & Tsitsa (1988) also examined instances of both stressed and unstressed Greek vowels (produced by 7 male and 7 female speakers) and found no overlap between articulatory spaces of vowels, suggesting that Greek vowels are distinct.

Vowel duration appears to be an interesting feature in the Greek[9] language. Although vowel durations vary naturally, it is not a feature that is phonemically relevant (Fourakis *et al*., 1999). There are no minimal pairs acoustically different by one phoneme differentiated by this dimension (however, minimal pairs that are orthographically different do exist (e.g. '*σήκω*', /siko/, stand up, vs. '*σύκο*', /siko/, fig)[10]. Previous studies attribute the variation of Greek vowel duration on the effects of speaking rate, stress, word length and prosodic context; for example it has been observed that generally vowels shorten when there is increase in the speaking rate (Fourakis *et al*. 1999), and stress has been observed to be a main aspect for vowel duration (Botinis, 1989; Arvaniti, 1991, 2000; Fourakis *et al*., 1999; Nicolaidis & Rispoli, 2005; Nicolaidis & Sfakianaki, 2007; Baltazani, 2007). Table 1.4 presents duration rates (in ms) of

[9] Ancient (Attic) Greek phonemically contrasted long and short vowels and the vowel inventory of Ancient Greek contained five short and seven long vowels that were realized as distinct phonemes (Allen, 1999). These features of Ancient Greek vowels are not present in Modern Greek. Only some orthographic representations have been preserved.

[10] This issue is also discussed in relation to reading processes and access to phonology in Experiment 5 (Chapter 5).

Greek vowels (values adopted from Arvaniti, 2000; Fourakis *et al*., 1999; & Nicolaidis, 2003). Arvaniti (2000) and Fourakis *et al*. (1999) make the distinction between stressed and unstressed vowels.

Table 1.4: Duration values (in ms) of Greek vowels adopted from Arvaniti (2000), Fourakis *et al*. (1999), Nicolaidis (2003). Arvaniti (2000), Fourakis *et al*. (1999) show values of vowel duration in stressed and unstressed position.

Vowel	Stress	Arvaniti (2000)	Fourakis *et al*. (1999)	Nicolaidis (2003)
i	stressed	106	76	69
	unstressed	77	44	
e	stressed	113	94	81
	unstressed	85	57	
a	stressed	126	105	85
	unstressed	89	78	
o	stressed	123	94	78
	unstressed	96	67	
u	stressed	120	88	60
	unstressed	89	54	

According to Fourakis *et al*. (1999), vowel duration is reduced by 40% in the case of unstressed vowels. These findings provide evidence that duration is longer for stressed vowels. Duration variation for stressed and unstressed vowels has been considered as context dependent[11]; it depends on the position of the vowel in the word and is closely related to stressed versus unstressed syllables, i.e. vowels in stressed syllables appear slightly longer than vowels in unstressed syllables (Fourakis *et al*. 1999; Mackridge, 1985). This effect, however, does not lead to a phonemic distinction between long and short vowels (similar to the case of English vowels discussed in section 1.5.2). Thus,

[11] Arvaniti (1991, 2000), Fourakis (1986) and Baltazani (2007) give detailed accounts with regard to how contextual variation affects vowel duration changes, however this is not discussed in detail here as it is beyond the scope of this study.

vowel length is related to syllable stress (or emphasis) placed on a vowel rather than vowel category classification. The same vowel can occur in a stressed and in an unstressed syllable, but in the case of stressed syllable it could be longer. Duration is also affected by word length; vowels in shorter words appear longer than vowels in longer words (Baltazani, 2007). This effect has also been observed in other languages such as English (Nakatani, O'Connor, & Aston 1981) and Japanese (Beckman, 1982).

1.5.2 English Vowels

English has a more complicated and larger inventory of vowels compared to Greek. English[12] vowels have variations on two dimensions, which co-vary simultaneously: *duration* and *spectral changes*, although it is very common in the literature for English vowels being grouped according to their durational features, long versus short, for example, there is a long /i:/ and a short /ɪ/, or a long /ɔː/ and a short /o/ (e.g. Roach, 1991). For long vowels usually the length mark '**:**' is used (the development of which is credited to Gimson, 1962), however it could be argued that this symbol is in fact redundant since the actual vowel symbols used are distinct (depicting on the relevant spectral location for vowels) and can thus successfully distinguish the vowels from one another (e.g. Ladefoged, 2001). The length mark symbol is however used especially when there is a need to emphasise the duration features of the vowels.

In Standard British English there are 11 monophthong vowel phonemes (Hawkins & Midgley, 2005; Wells, 1962): /a/, /ɑ/, /ʌ/, /ʊ/, /u/, /ɒ/, /ɔ/, /e/, /ɜ/, /ɪ/, /i/ (see Figure 1.6). There is also the schwa /ə/ vowel which, although it has no unique orthographic representation, exists as a weakly pronounced central vowel. However, due to the fact that the schwa /ə/ does not appear in stressed syllables, it is not discussed further in this thesis (although it is included in some figures and tables for a complete and comprehensive account of the English vowel inventory).

[12] English vowels throughout this thesis always refer to Standard British English or else RP (received pronunciation) non-accented English.

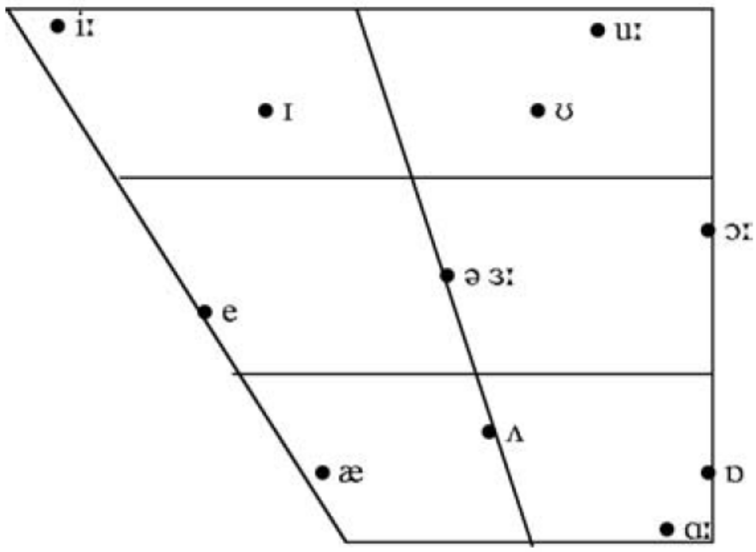


Figure 1.6: Vowel monophthongs of the English vowel inventory.

Figure 1.6 presents the distribution of English vowels in terms of their place of articulation and duration contrasts, i.e. high-mid-low, front-central-back, long-short[13]. Table 1.5 presents an overall synopsis of the English monophthong vowels and their distribution in the relevant categories and Figure 1.7 graphically shows the spectral characteristics of the English vowels which reveals a correlation with their respective place of articulation.

Table 1.5: Height and horizontal articulatory positions of British English vowels.

	Front		Central		Back	
	long	short	long	short	long	short
High (or Close)	i:	ɪ			u:	ʊ
Mid		e	ɜ:	ə	ɔ:	
Low (or Open)		a		ʌ	ɑ:	ɒ

[13] The length mark ':' is used to signify long vowels.

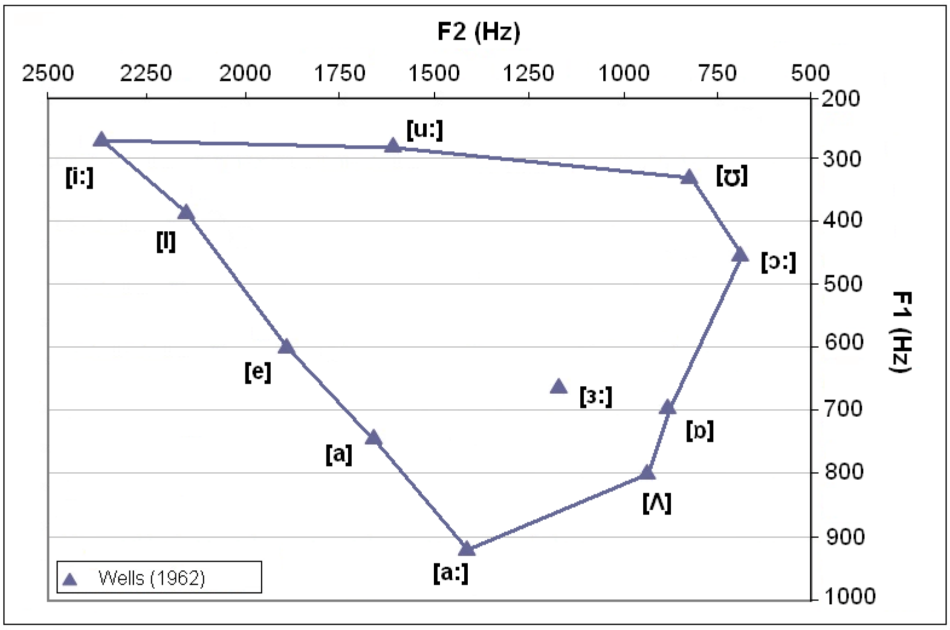


Figure 1.7: Graphic representation of mean F1 and F2 values for English vowels (mean values adopted from Wells, 1962).

Hawkins & Midgley (2005) report typical formant frequency values (F_1 and F_2) for the British English vowels. They tested four different age groups (20-25, 35-40, 50-55, 65+) but only the 20-25 age group is considered here (Table 1.6) so that they can be compared with the formant values produced by Greek adult speakers of the same age group (20-28 years).

Table 1.6: Mean formant values F_1, F_2 (in Hz) of the British English vowels produced by English adult speakers (values adopted from Hawkins & Midgley, 2005).

	a	ɑ	ʌ	ʊ	u	ɒ	ɔ	e	ɜ	ɪ	i
F_1	917	604	658	413	289	484	392	600	494	393	276
F_2	1473	1040	1208	1285	1616	865	630	1914	1373	2174	2338

In English, both spectral and durational features are naturally present and contribute to vowel quality. For instance, the /i/ versus /ɪ/ vowel production would not only reveal differences in the formant values of each segment (as presented in Table 1.6) but would also have distinct duration measurements. Mean duration values for all British English vowels are presented in Table 1.7.

Table 1.7: Duration values (in ms) of British English long and short vowels adopted from Wells (1962).

Vowel	Length	Wells 1962
a	short	210
ɑ	long	335
ʌ	short	148
ʊ	short	142
u	long	294
ɒ	short	178
ɔ	long	330
e	short	170
ɜ	long	309
ɪ	short	139
i	long	293

Minimal pairs in English form a very appropriate context for acoustic features of English vowels to be considered. For instance, the two respective vowels in the minimal pair /si:t/ (seat) - /sɪt/ (sit) would involve the following acoustic characteristics: spectrally, the /i:/ would be a high-front vowel (F_1 276Hz x F_2 2338Hz), the /ɪ/ would be a high-front vowel (F_1 393Hz x F_2 2174Hz); duration values would be 293ms and 139ms respectively. Numerical values show that even in the case of two vowels being located spectrally within the same category (e.g. high-front) there are still individual features that contribute toward the categorization of vowels into distinct categories[14].

[14] More specifically, the extent to which the removal of one element (e.g. duration) may pose an effect on vowel perception and categorization (by native and non-native speakers alike) is one of the main objectives of this thesis and is considered in later chapters.

1.5.3 Comparison of the Greek and English Vowel Inventories

The vowel inventories of Greek and English differ in several important ways. First, in the number of vowel phonemes involved in each of the two vowel inventories. In Greek there are 5 monophthong vowels [i, e, a, o, u], in English there are 11 monophthong vowels [/a/, /ɑ/, /ʌ/, /ʊ/, /u/, /ɒ/, /ɔ/, /e/, /ɜ/, /ɪ/, /i/].

Second, duration is a feature that naturally occurs in both languages; however its use is not the same in Greek and English. As already discussed in previous sections, duration in Greek is found as an additional (or suprasegmental) acoustic feature, not directly characterizing vowel category; stressed vowels are longer and this depends on the lexical context as opposed to the same vowel being in unstressed lexical context (and therefore relatively shorter) (e.g. Fourakis *et al*., 1999). Duration in English however is differently used as every English vowel carries unique duration values and this does not relate to stressed versus unstressed conditions as in the case of Greek (Table 1.8 presents an overall account for the mean duration values for Greek and English vowels).

Table 1.8: Duration values (in ms) of Greek vowels (adopted from Arvaniti, 2000; Fourakis *et al*., 1999; Nicolaidis, 2003) and English vowels (adopted from Wells, 1962).

Greek Vowel	Stress	Arvaniti 2000	Fourakis et al. 1999	Nicolaidis 2003	English Vowel	Length	Wells 1962
i	stressed	106	76	69	**i**	long	293
	unstressed	77	44		**ɪ**	short	139
e	stressed	113	94	81	**e**	short	170
	unstressed	85	57		**ɜ**	long	309
a	stressed	126	105	85	**a**	short	210
	unstressed	89	78		**ɑ**	long	335
ɒ	stressed	123	94	78	**ɒ**	short	178
	unstressed	96	67		**ɔ**	long	330
					ʌ	short	148
u	stressed	120	88	60	**u**	long	294
	unstressed	89	54		**ʊ**	short	142

Thus, although there are no durational distinctions between Greek vowels as in the case of English vowels (long versus short), slight duration changes occur in the case of stressed vowels yet duration is not a distinctive feature for vowel classification. English vowels, however, can be grouped in two main groups based on duration values, i.e. long vowels or short vowels. Overall, vowel duration in Greek does not play a part in phonological contrast, it occurs but not contrastively. In English, vowel duration can be a cue for vowel categorization however other cues (such as spectral cues) can be critical for phonological contrast[15]. English vowels are often categorised as (relatively) short or (relatively) long in the literature (e.g. Roach, 1991), however duration is not the only separable difference between them, unlike other languages, for example Finnish where duration is phonemically relevant; duration is used phonologically in quantity distinctions and for Finnish vowels duration change is the sole dimension that differentiates two –spectrally identical- vowels (e.g. Wiik, 1965; Ylinen *et al*., 2010).

Botinis, Bannert, Fourakis, & Pagoni-Tetlow (2002) examined the effects of suprasegmental features (such as, stress, syllable position, focus, and tempo) on segmental durations in British English and Greek. They found that stress has significant effects on vowel duration for both British English and Greek. As Botinis *et al*. (2002) claim, these results are suggestive and not conclusive due to a number of reasons (e.g. data quantity poses statistical shortcomings, the use of non-words rather than words as stimuli) and further investigations are required in order to shed more light and evaluate the above assumptions.

Third, all Greek vowels have equivalent counterparts (or, are 'shared') with the English vowel inventory, but in English there are vowel categories that are not present in Greek. Attempting to pair those 'shared' (or native) vowels (Giegerich, 2001) with those that could be considered as new (non-native) entities for a Greek speaker of L2 English, a remarkable symmetry can be observed: /a/-/ɑ/, /a/- /ʌ/, /ʊ/-/u/, /ɒ/-/ɔ/, /ɛ/-/ɜ/, /ɪ/-/i/; the first counterpart of each pair is the 'shared' vowel between Greek and English, whereas the second

[15] This is experimentally explored in later chapters.

counterpart of each pair is a non-shared entity with respect to the Greek system of vowel phonemes. Therefore, the /a/, /ʊ/, /ɒ/, /ɛ/ and /ɪ/ vowel phonemes are considered as 'shared' (or native), while the /ɑ/, /ʌ/, /u/, /ɔ/, /ɜ/ and /i/ vowel phonemes are considered as new (or non-native) (Flege, 1988) (see Table 1.9).

Table 1.9: Mean formant values F_1, F_2 (in Hz) of the Greek and British English vowels (adopted from Sfakianaki, 2002 and Hawkins & Midgley, 2005).

Greek Vowel	F1	F2	English Vowel	F1	F2
i	423	2073	**i**	276	2338
			ɪ	393	2174
e	601	1811	**e**	600	1914
			ɜ	494	1373
a	736	1466	**a**	917	1473
			ɑ	604	1040
ɒ	434	921	**ɒ**	484	865
			ɔ	392	630
			ʌ	658	1208
u	583	1137	**u**	289	1616
			ʊ	413	1285

Considering the study by Bradlow (1995) who compared four 'shared' vowels [i, e, o, u] between American English, Spanish and Greek as previously discussed, there were differences in the F_2 values between the three languages. The term 'shared', therefore, cannot be considered an accurate description but is rather vague. Looking at spectral differences between the Greek and English vowels, the mean formant values compared in Figure 1.8 demonstrate language-specific differences; plotting the F_1 x F_2 mean values of the Greek and English 'shared' vowels, it becomes apparent that each of these 'shared' vowel categories are spectrally in close approximation but do not precisely match.

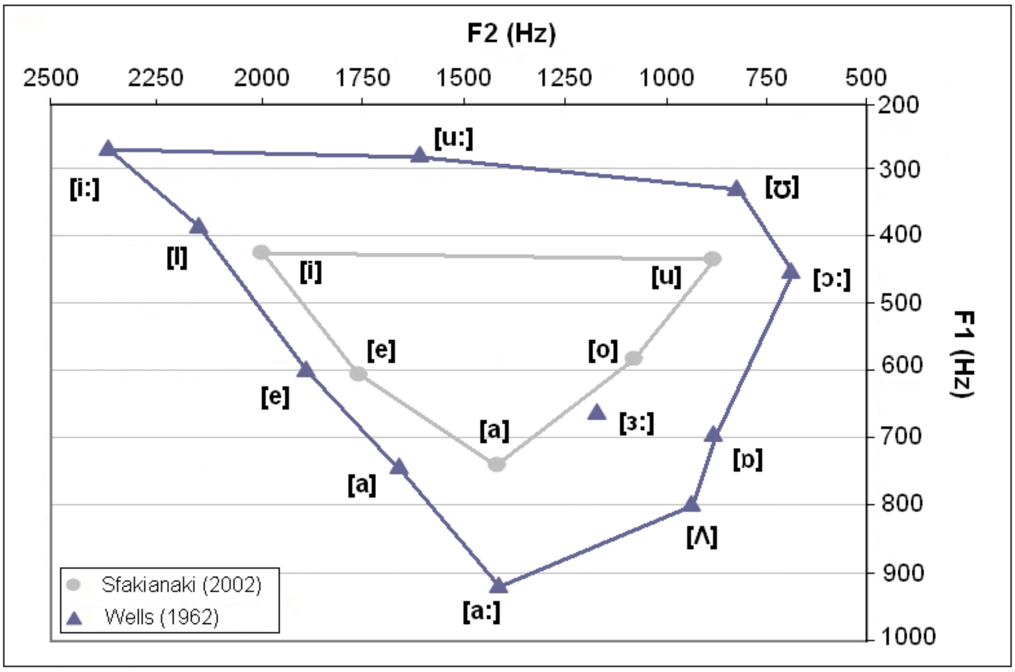


Figure 1.8: Graphic representation of mean F1 and F2 values for Greek and English vowels (mean values adopted from Sfakianaki (2002) and Wells (1962) respectively).

Greek and English vowels differ in the position occupied by phonemes 'shared' by both languages in the vowel space (see Figure 1.8). In general, English vowels are more peripheral than Greek vowels; in other words, English vowels are produced closer to the edges of the vowel space (see Figure 1.8).

In a study by Hawks & Fourakis (1995), Greek and American participants were tested in a perceptual identification task using synthetic vowel stimuli. It was observed that Greek participants separated the acoustic space into subspaces, equivalent to the number of Greek vowel categories, and also there was overlap of the American vowel distinction [e] – [ε] with the Greek [e] vowel category and the [o] – [ɔ] distinction with the Greek [o] vowel category. This indicates that the native inventory of speakers can affect perception and categorization of speech sound categories and classify 'similarly' sounding phonemes into native categories rather than assign them new perceptual subcategories. This issue is discussed in more detail in later chapters.

Second language learning involves not only the learning of L2 speech sounds but essentially includes learning of the relation between those speech sounds and how these are orthographically represented. This process gradually leads to the development of L2 reading skills. The following section explores the issue of reading processes and specifically compares the orthographic systems of English and Greek, pointing out their different levels of transparency which could pose a stumbling block for the L2 learner.

1.6 Orthography: from phonological development to reading

Following early stages of phonological development, and at around the (pre-school) age of five, children start learning to read. Reading is a skill that is developed over years of practice throughout one's school years and beyond, as one of the cornerstone skills for modern day language acquisition. Reading enables access to meaning through the mapping between the visual symbols used in a given language and the sound units that consist the phonology of that language (e.g. Ziegler & Goswami, 2006). Reading could thus be described as a process of 'phonological recoding' (Share, 1995, p.152). Every skilled reader knows that reading is not just a simple process but a quite complex one that essentially takes place between the visual identification of letter-strings and comprehension of words.

Word length, word frequency, word familiarity, word regularity, all contribute to the speed of reading (e.g. Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Balota & Chumbley, 1984). In terms of word length, short words are generally read faster than longer words (e.g. Frederiksen & Kroll, 1976; Balota & Chumbley, 1985; Wydell, Vuorinen, Helenius, & Salmelin, 2003; Weekes, 1997; Juphard, Cabonnel & Valdois, 2004). Further, the length effects are modulated by word frequency particularly in English. That is, low frequency words show larger length effects as compared with high frequency words (e.g. Content & Peereman, 1992); high frequency words are also reported to show no length effects (e.g. Weekes, 1997). In addition non-words (made-up words) in general yield longer reading times (RTs) than words, and this is called a lexicality effect (e.g. Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Lukatela, Popadic, Ognjenovic, & Turvey, 1989; Ziegler, Perry, Jacobs, & Braun, 2001; de Groot, Borgwaldt, Bos, & Van den Eijnden, 2002; Wydell *et al*., 2003). The regularity or irregularity of word spelling (i.e. regular word spelling would follow grapheme-to-phoneme conversion (GPC) rules while irregular word spelling does not) can also affect RTs with readers, in languages such as English or French (see section 1.6.1 on orthographic transparency for different languages), often producing longer RTs in reading words with irregular spelling leading to regularity effects (e.g. Glushko, 1979; Andrews, 1982; Monsell, Patterson, Graham, Hughes, & Milroy, 1992). Similar to the length effects, the regularity effects are also modulated by the word frequency, that is, the regularity effects are attenuated for high-frequency words (e.g. Ziegler *et al*., 2001; de Groot *et al*., 2002). The spelling of the word also affects the speed in which a word (spoken or read) is recognized, known as the *orthographic consistency effect* (e.g. Ziegler & Ferrand, 1998; Pattamadilok, Morais, Ventura, & Kolinsky, 2007; Ziegler, Petrova, & Ferrand, 2008; Pattamadilok, Morais, De Vylder, Ventura, & Kolinsky, 2009; Peereman, Dufour, & Burt, 2009). That is, word pairs that rhyme with only one possible spelling (e.g. *team* - *cream*) are read faster than word pairs that rhyme and can have different spellings (in terms of the word-final letter clusters or letter combinations) (e.g. *crane* - *train*).

1.6.1 Orthographic transparency & Granularity problem

The graphic representation of sounds that combine to make words is the orthographic system of a language. Different languages use different systems, such as alphabetic systems (where written symbols /graphemes and spoken phonemes are mapped, e.g. latin-based languages), logographic systems (where symbols represent words, e.g. Chinese or Japanese Kanji), or syllabic systems (e.g. Japanese Kana or Korean Hangul). A grapheme is the basic written form of a structural unit in any given language, representing a speech sound (or phoneme) and can be in the form of alphabetic letters, characters (e.g. Chinese or Japanese), numerical digits, punctuation marks, or individual symbols used by the different writing systems of the world (e.g. Goswami, 2008). In English, for example, graphemes are the individual written letters in the alphabet.

The grapheme-phoneme correspondence however, is not always a 1:1 mapping particularly in English. The degree of grapheme-to-phoneme mapping for each language has been described as *orthographic transparency* (e.g. Ellis, 2001; Ellis & Hooper, 2002; Ward, 2010; Wydell & Butterworth, 1999). Languages could be viewed as forming a continuum, with various levels of orthographic transparency; languages that follow a 1:1 grapheme-to-phoneme mapping are often described as having transparent or shallow orthographies (e.g. Italian, Spanish, Finnish, Welsh, Dutch, Turkish, Greek or Japanese Kana) while the extreme opposite orthographic mapping is described as opaque or deep orthography (e.g. English, French, Hebrew or Japanese Kana). Learning to read a language with transparent orthography is thus easier than a language with opaque orthography (Seymour, Aro, & Erskine, 2003).

Some languages include multiple orthographic units which correspond to single phonemes, for example, the English phoneme /f/ can be represented as 'ph' as well as 'f'. This is the so-called 'granularity problem' (Ziegler & Goswami, 2005; Goswami, 2008). Becoming a proficient reader in such a language could be suggested to be a harder task. For example, children who learn how to read in English are required to develop whole-word naming strategies for reading words such as 'yacht' or 'tough', also develop word rhyming strategies for reading words such as 'might', 'right', or 'light', and grapheme-phoneme conversion strategies are required for words such as 'mat', 'mug', or 'bid'[16] (e.g. Goswami, 2008).

Children show remarkable reading skills even from the first few months of reading instruction for languages such as German, Italian, Turkish, and Finnish; they tend to develop a grapheme-phoneme strategy which is sufficient

---

[16] These examples, of course, do not exclude the use of rhyming strategies, e.g. '*mat*' rhymes with '*pat*' or '*rat*' or '*bat*' or '*cat*'.

in accurate reading in such languages as experimental studies have shown (e.g. Cossu, Gugliotta, & Marshall, 1995; Wimmer, 1996; Durgunoglou & Oney, 1999; Goswami, 2008). It is possible to test children's reliance on grapheme-to-phoneme conversion strategies by looking at length effects, i.e. the more letters a word contains the longer it should take to read. This is true in the case of languages with transparent orthographies such as Greek, with RT results on word length effects being more consistent, as opposed to data from English readers (Goswami, Porpodas, & Wheelwright, 1997).

Non-word reading is also another interesting area as there are two types of non-words: those whose spelling corresponds to real words and those whose spelling does not correspond to real words. Therefore it is possible to read a non-word with spelling that corresponds to a real word through applying analogy (Goswami, 2008). However, if applying grapheme-to-phoneme conversion strategies, there should be no reaction time difference in reading either type of non-words. It has been suggested however that results of reading non-words in languages such as German that employs a rather consistent orthography are more accurate compared to the equivalent level of English readers (Frith, Wimmer, & Landerl, 1998). In a study by Goswami, Ziegler, Dalton, & Schneider (2003), English readers were more accurate at reading non-words that correspond to real English words that can be read by analogy, as opposed to reading non-words that do not have English real word correspondents. German children, however, showed no accuracy differences in reading either non-word types.

Seymour *et al*. (2003) conducted a large cross-language study that compared first-grade child readers in 14 European languages. Readers were tested for words and non-words. All words were matched for difficulty across all participating languages, and length of reading instruction was also equated. However difference in the age when children start attending first-grade at school would vary between languages due to the difference in the respective educational systems. Results showed that for languages such as Greek, Finnish, German, and Spanish, accuracy levels for words and non-words were between 90-98%. In the case of languages with inconsistent (or opaque) orthographies lower accuracy levels were observed, e.g. French (79%), Danish (71%), and English yielded lowest scores with word accuracy at 34%. These results emphasize the effect of orthographic transparency and also suggest that English is a rather complicated orthographic system to master compared to other European orthographic systems (e.g. Greek).

1.6.2 Reading Processes and Cognitive Models of Reading

A vast amount of research on reading processes has been conducted using cases of reading impairments such as developmental or acquired dyslexia[17]. The observation of different reading impairments can provide insights into the complex non-impaired reading process (e.g. Marshall & Newcombe, 1973). For instance, surface dyslexic patients tend to regularize irregularly spelt words, such as reading the word 'pint' as if rhyming with 'hint', 'mint' or 'tint', suggesting that their lexical reading route is thought to be impaired. Patients with phonological dyslexia cannot read non-words or novel words that are not stored in their mental lexicon but demonstrate ability to read familiar words. They usually tend to read non-words as visually similar (and similarly sounding) familiar words, e.g. '*cobe'* as '*comb'*, *'ploon'* as '*spoon'*, *'fude'* as '*fudge'* (Funnell, 1983). Individuals with phonological dyslexia read on a whole-word basis and are impaired at the sub-lexical processing level[18] which is required for phonological recoding in order to read unfamiliar words and non-words. Both lexical and sub-lexical reading impairments are associated with deep dyslexia. Additionally, deep dyslexic patients also tend to make semantic errors, e.g. reading the word '*sick*' as '*ill*', '*cheer*' as '*laugh*', or visual and then semantic errors, e.g. reading the word '*sympathy*' as '*orchestra*' via 'symphony' (Marshall & Newcombe, 1973; for a review see, Gazzaniga, Ivry, & Mangun, 1998; Price, 2000; Price & Mechelli, 2005).

---

[17] It is not the purpose of this section to exhaust the topic of reading impairments, however a short review is provided in order to outline the background that leads to theoretical models and frameworks.
[18] Sections 1.6.2.1 and 1.6.2.2 provide more information on the various routes involved in the reading process.

Reading is thus not just a mere conversion of written text to spoken language but involves orthographic processing, semantic processing, phonological decoding, lexical access, employing retrieval of meaning, requiring the contribution of memory, and more. In other words, it is a highly cognitive process. The brain, however, appears ready to 'cope' with such high volume cognitive task. For example, children develop reading skills in their native language, even becoming mature readers by the age of eight[19] (Paulesu, 2006). Or, even becoming competent readers in additional languages, which nowadays is not the exception but a common trend (e.g. Brizic, 2006; Jessner, 2008; Jessner, 2006; Cenoz & Jessner, 2000; Edwards, 1994).

The combination of different processes involved in reading and insights from the various reading impairments, as already discussed, suggest that separate mechanisms may be involved in mediating all different functions in normal reading performance. A number of theories, frameworks and computer-simulated models have emerged, attempting to explain how these processes function during normal (i.e. non-impaired) reading. Here the two most prominent models, namely '*Dual-Route Cascade reading model*' (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Coltheart & Rastle, 1994; Rastle & Coltheart, 1998) and '*Parallel Distributed Connectionist reading model'* (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996; Plaut & Kello, 1999), are presented.

1.6.2.1 Dual-Route Cascade Reading Model

The '*Dual-Route Cascade model*' (Coltheart *et al.*, 2001; previous version appeared in Coltheart & Rastle, 1994; Rastle & Coltheart, 1998), or DRC model, is a computer simulation model of reading and aims to explain the process of reading by skilled readers. It consists of two mechanisms, a lexical (semantic and non-semantic direct) route and a sublexical /phonological route (also referred to as Grapheme-to Phoneme Conversion (GPC) route). Each

[19] Specifically when learning a transparent orthography language.

route consists of a number of layers which interact. Each layer contains a set of units that correspond to the smallest parts of the model, for example, words in the orthographic lexicon and the letter unit layer respectively (see Figure 1.9). It is suggested that the units of the different layers interact in two ways: through inhibition or excitation. Inhibition refers to the case when the activation of a unit makes it harder for other units to be activated and excitation refers to the case when the activation of a unit contributes to the activation of other units (Coltheart *et al.*, 2001). Also, units on the same level may interact by lateral inhibition (this is depicted in Figure 1.9 with the arrows that aim to show the interactions between the various units).

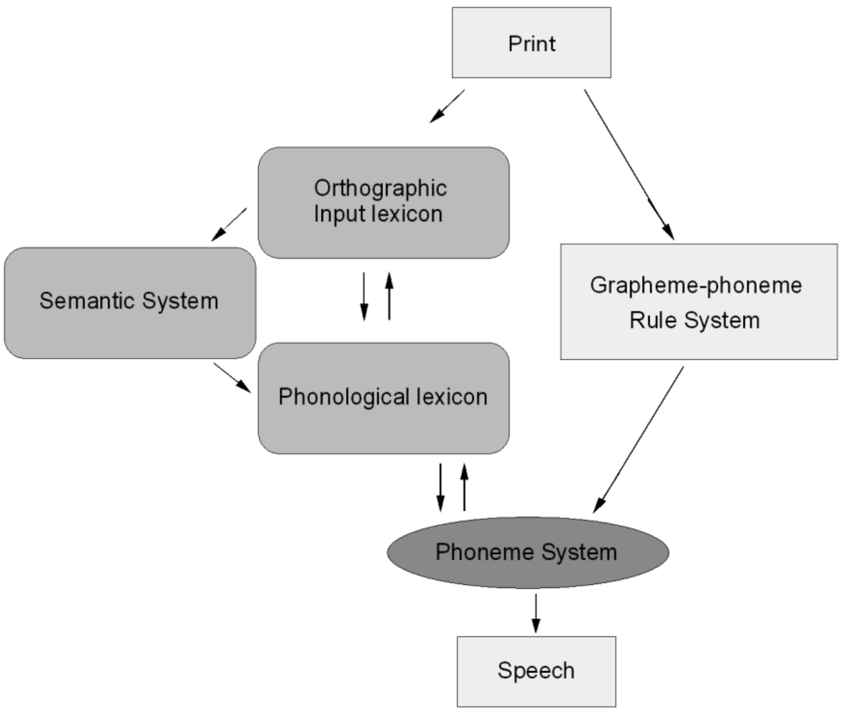


Figure 1.9: Dual-Route Cascade Reading Model (after Coltheart *et al.*, 2001; Coltheart & Rastle, 1994; Rastle & Coltheart, 1998).

The lexical route is a quick process where the orthographic form of the whole word is accessed in the orthographic input lexicon. Then, it either goes to the phonological lexicon directly, which accesses the whole word's phonology or it goes through the semantic system to the phonological lexicon, which requires the mediation of semantic knowledge of stored semantic information (e.g. concepts). The sublexical /phonological route is a sequential process, whereby graphemic information is serially transformed into phonology, therefore it is a grapheme-to-phoneme conversion (GPC) process which bypasses the orthographic lexicon, the semantic system, and phonological lexicon. The GPC route enables the reading of novel words or non-words which do not have stored lexical entries that may be retrieved. Both lexical and sub-lexical routes converge at the 'phoneme system' level which functions as the preparation stage of the phonological output for articulation.

The DRC model proposes that during natural reading both routes are activated simultaneously and competitively as to which route would generate the correct response for a given text. Reaction time for word naming is expected to be faster for high-frequency words when the lexical route is activated. However, for low-frequency words, semantic access is also required, hence longer reaction time is expected. At the same time, the sub-lexical phonological route could be activated. This could initiate a 'conflict' between the lexical and sub-lexical routes.

The contribution of other factors such as word regularity or word length may also be contributing factors for word naming. For example, in the case of low-frequency irregular words as opposed to high-frequency irregular words, a 'lexical lookup procedure' (Rastle & Coltheart, 2000, p. 343) is required for correct articulation and the 'conflict' between the lexical and sub-lexical routes leads to longer reaction times. In the case of novel or non-words sublexical /phonological (GPC) recoding is required. However, certain non-words could activate the GPC route and then activate a lexical lookup which would generate possible rhyming options. For example, the non-word 'vint' would require GPC processing followed by lexical lookup which would determine whether it could rhyme with either 'pint' or 'mint' (e.g. Rastle & Coltheart, 2000). Word-length may also affect the amount of time required for processing in the case of words and non-words alike (e.g. Rastle & Coltheart, 1998; Weekes, 1997).

1.6.2.2 Parallel Distributed Connectionist reading model

The Parallel Distributed Connectionist (PDP) reading model (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996; Plaut & Kello, 1999) is an alternative to the DRC reading model and proposes that there is a single mechanism that processes both words and non-words. According to the PDP model three main units, i.e. orthography, phonology, and semantics, interact in order for a correct phonological sequence to be generated following the orthographic input. Existing internal representations distributed within this network, referred to as 'hidden units' (Seidenberg & McClelland, 1989: 526; see Figure 1.10), also contribute to this 'weighting' among the three main units, and are modulated based on the reader's experience with the orthographic patterns (i.e. visual exposure) involved in the print input. The three main units form a network mediated by the hidden units, work interactively during the process of reading, and consist of two layers: visual input and articulatory output.

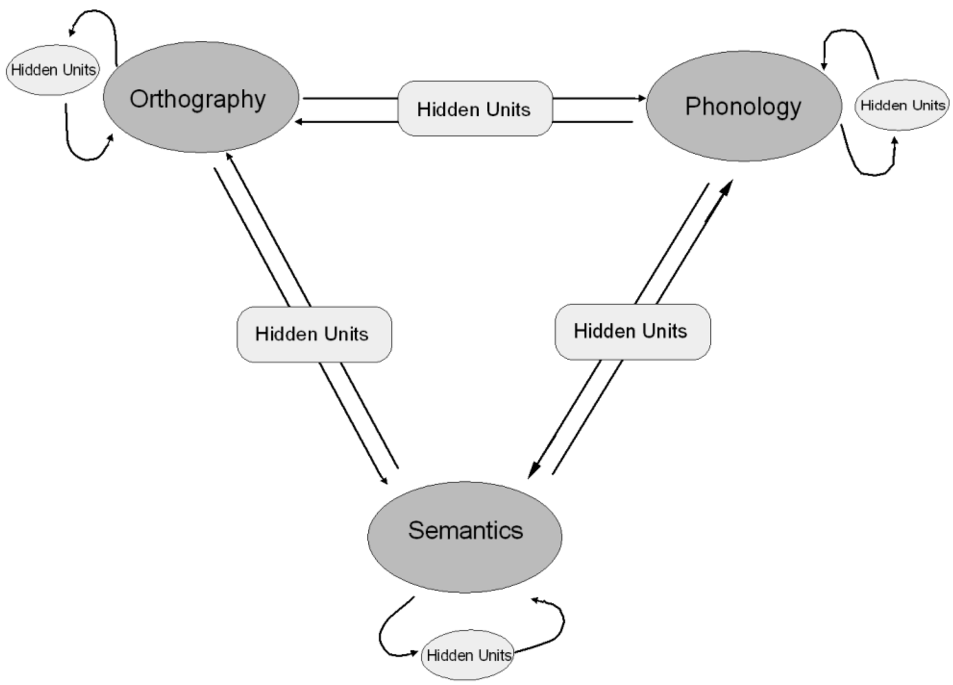


Figure 1.10: Schematic representation of the Parallel Distributed Connectionist reading model (after Seidenberg & McClelland, 1989).

Similar to the DRC model, the degree of word familiarity, regularity, lexicality and frequency affect the interaction of units that are activated during the reading process. When a familiar word is read, the model generates a representation of the entire word while semantic weighting also takes place prior to articulatory output. The regularity or irregularity of the word also determines the amount of semantic weighting required, triggering longer reaction time if the word is irregular (Plaut *et al.*, 1996). In the case of a non-familiar (novel) word, a sequence of phonemes is activated and grapheme-position is also detected. Visual fixation occurs until the graphemic and phonemic units are monitored as regular correspondents. However, numerous visual fixations on the visual stimulus are required if the word is an irregular novel word or a non-word. In this case, a phoneme-by-phoneme detection occurs until an acceptable pronunciation becomes possible. The exact way that these procedures are activated, e.g. in parallel or sequentially, is still debatable (e.g. Foorman, 1994). Various experimental studies however have supported their existence through word naming tasks measuring factors such as reaction time or accuracy levels.

The PDP model represents a set of computational principles that aim to resemble properties of neural computation. Practical applications to reading processes focus on the learning of the internal distributed representations that mediate the interaction of orthographic, phonological and semantic information. This computational system described in the PDP model, however, lacks word-specific representations and could be considered as limited in the capacity and diversity of vocabulary it can handle (Plaut, 2005), which 'weakens' the 'semantic unit' module. This indicates that the PDP stands in sharp contrast with the Dual-Route Cascade (DRC) model in their implications for word reading. The DRC seems more applicable to a wide variety of empirical issues (e.g. the reading of non-words), compared with the PDP that appears more limited in the empirical issues it can address. It could be suggested that both models could provide insights into reading processes, but their implications and limitations are further addressed in Chapter 5 in view of experimental data.

1.6.3 Greek vs. English orthography

As already discussed, every language has a different level of orthographic transparency. In this respect, Greek orthography is very different to English orthography. This section aims to look into the orthographic systems of the two languages and compare the two systems.

Greek is a consistent and transparent language for reading where there is a regular grapheme–phoneme correspondence. Two main factors contribute to this, consistency (consistent mapping of graphemes and phonemes) or regularity and structural simplicity (also reflected in articulation of Greek words). Greek graphemes are represented by one phoneme for most Greek words and there are no words with exceptional spellings unlike English (e.g. thought, yacht, light). There is also consistency in vowel pronunciation unlike English vowels that can be shaped according to neighboring letters or be pronounced as schwas /ə̲/ (as in 'potato' /pə̲teɪtə̲/). Consonants are always pronounced, in contrast to English which includes words where consonants exist in the spelling but are totally disregarded when the word is pronounced (e.g. island, walk, psychology). The reliable mapping between graphemes and phonemes that exists in Greek makes reading easier and reduces the need to learn whole-word pronunciations (Nikolopoulos, Goulandris, & Snowling, 2003). Also, the open consonant-vowel (CV) structure of most cases of Greek syllables is an additional feature of Greek as a transparent reading system, and more specifically indicates phonological transparency (Nikolopoulos *et al*., 2003). Contrastively, many English words involve sequences of consonant clusters (e.g. thought, kitchen) which render the phonological structure of English and consequently reading in English is much more complicated in comparison with Greek.

Although the regularities in grapheme-phoneme mappings characterize Greek as a transparent language for reading, spelling is a feature that is less transparent, or at least involves some ambiguities and this is mainly caused by the variety of grapheme representations for the Greek vowels (Nikolopoulos *et al*., 2003). For example, the phoneme /ɒ/ can be represented by two different graphemes, as ‘ο’ or ‘ω’, the phoneme /ε/ can also be represented by two different graphemes, ‘ε’ or ‘αι’. Similarly, the phoneme /iː/ can be represented by 6 different graphemes, ‘ι’, ‘η’, ‘υ’, ‘ει’, ‘οι’, ‘υι’ and interestingly different representations of a single phoneme can even occur within a single word (e.g. *κυνήγι* /kiːˈniːɣiː/ = *hunting*, where the /iː/ sound is spelt with three different grapheme representations in this one word). These spelling patterns in Greek reflect historical changes that have occurred and have affected the spoken as opposed to the written language. In Ancient Greek (4th century BC) differences in the spelling of a certain vowel sound would indicate pitch alterations which are no longer used in Modern Greek speech (Nikolopoulos *et al*., 2003). This makes it possible for a Modern Greek word to be pronounced correctly even in the case that the spelling is incorrect (e.g. *κυνήγι* /kiːˈniːɣiː/ = correct Greek spelling for ‘*hunting’* vs. *κινήγυ* /kiːˈniːɣiː/ = incorrect Greek spelling for *‘hunting’*). In English, although there are cases of two different spellings for a single phoneme sound (e.g. /f/ is spelt as ‘f’ and ‘ph’ or /i/ as ‘i’ in *bit* or as ‘e’ in *before*), English phonemes are not grammatically realized in the same way as in Greek (as in the example described above).

Spelling in Greek also reveals that it is a highly inflected language system. Inflectional morphemes are used in order to mark tense (present, past, future, etc.), case (nominative, accusative, genitive, dative, vocative), number (singular or plural), person (I, you, we), mood (indicative, subjunctive) and voice (active or passive; for a review see Giannakopoulou, 2004). For example, first, second and third person singular and plural in verbs are indicated by different endings which are also influenced by the grammatical status of the verb as different verb type requires different affixes (endings) (e.g. κάνω /kanɒ/ =I do, κάνεις /kɒniːs/ =you do, κάνει /kaniː/ =s/he/it does, κάνουμε /kanume/ =we do, κάνετε /kanete/ =you do, κάνουν /kanun/ =they do versus σκέφτομαι /skeftome/ =I think, σκέφτεσαι /skeftese/ =you think, σκέφτεται /skeftete/ =s/he/it thinks, σκεφτόμαστε /skeftɒmaste/ =we think, σκέφτεστε /skefteste/ =you think, σκέφτονται /skeftɒde/ =they think). English, however, makes use of few inflectional patterns, for example final ‘–s’ is used for verbs in the third person singular (e.g. s/he walks), or for nouns in order to mark plural (e.g. bat – bats), and '-ed' suffix is added to verbs[20] in order to mark past tense. It becomes obvious that in order for correct spelling in Greek, extensive knowledge of grammatical rules that govern inflectional patterns is required, but mastering reading in Greek is an easier skill to acquire compared with English that involves little inflectional patterns yet it employs a more opaque orthography.

Another interesting case is homophones in relation to orthography. For instance, in Greek there are words that are spelt differently but sound exactly the same only because a vowel phoneme is spelt using a different grapheme in each case (e.g. σήκω /ˈsiːkɒ/ =stand up, verb vs. σύκο /ˈsiːkɒ/ =fig, noun). A grapheme-to-phoneme conversion mechanism would enable a reader to read such words correctly (Loizidou-Ieridou, Masterson, & Hanley, 2010). In English, however, in order for correct homophone pronunciation to be achieved it is likely that whole-word phonological access may often be required (e.g. *blew – blue* as regular homophones, *I – eye* as irregular homophones).

Overall, the Greek orthographic system is a transparent system which is easy for readers to master accurately even from early stages of reading instruction (as the study by Seymour *et al*. (2003) mentioned in section 1.6.1 reveals), however spelling in Greek is less consistent. Grammar, morphology and syntax are interdependent and interactive factors in Greek printed text which affect the use of inflection patterns as well as meaning. The English orthography, however, is not transparent on the reading level despite the limited use of inflectional morphemes. This suggests that it is possible for native readers of a language to be affected in their reading performance when they become learners of a L2. This thesis considers this assumption within the more general aim of L2 learning, specifically looking at Greek learners of L2 English and how specific language patterns of either their L1 or the L2 may affect reading performance.

---

[20] The addition of the '-ed' suffix applies to regular verbs (e.g. play – play<u>ed</u>) but does not apply to irregular verbs (e.g. go - went) which are an exception.

1.7 This thesis: Research predictions and hypotheses

This thesis examined plasticity in second language learning. Specifically, native Greek speakers of L2 English were the main target group. The three main strands of this thesis were explored via five experimental studies.

The first strand explored cross language effects in speech perception and with particular emphasis on cue-weighting effects. The main research questions explored here were: whether L2 learners utilise the same or different acoustic cues for perception of L2 sounds as native speakers of the target language do, and also whether there is interference between the L1 and the L2 for L2 speech perception. Experiment 1 administered perceptual identification and discrimination tasks testing perception of the English tense-lax /iː/ vs. /ɪ/ vowel contrast. Child and adult participants were tested in order to also investigate age effects and the possibility of language experience (L1 vs. L2) being an important and contributing factor for L2 vowel identification and discrimination. Specific acoustic features (i.e. duration and spectral features /cues) of the English vowel contrast were manipulated (see Section 2.1) in order to investigate cue-weighting effects.

It is hypothesised here that Greek L2 learners of English could pay attention to acoustical cues that are not used by native speakers of English for correct identification of this vowel contrast (e.g. Holt & Lotto, 2006; Iverson *et al.*, 2003). L2 speech sounds are also expected to assimilate into a single L1 speech category as the PAM would also predict (Best, 1995a; Best & Tyler, 2007), particularly given that for the tense – lax (/iː/ vs. /ɪ/) English vowel contrast there is one equivalent 'i' category in the Greek vowel inventory (see section 1.5.3 for a discussion). Also, this study assesses the NLM proposal (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995) that the already established L1 speech sounds are powerful attractors to native categories that can impair the formation of new L2 categories.

Experiment 1 also discusses the possibility of L1 transfer in view of the underlying phonological systems (Greek vs. English). Two contrasting hypotheses are examined: 1. if a feature is not used contrastively in the L1 it is difficult for a category to be formed (McAllister, Flege, & Piske, 2002); or, 2. L2 learners can be sensitive to novel acoustic features even if not used contrastively in their L1 (Bohn, 1995). These hypotheses are discussed and evaluated in view of the experimental results. Age related effects are also expected, especially considering the degree of L2 experience for Greek adult L2 learners of English compared to child L2 learners (e.g. as predicted by the SLM, Flege, 1981, 1987, 1991a, 1992, 1995a). Adults are thus predicted to perform better than children.

To the same direction, Experiment 2 aimed at examining the existence of orthographic cues in the case of word pairs. The same research questions were explored as in Experiment 1. However, Experiment 2 employed a slightly different experimental design: it used picture representations of minimal word pairs instead of their orthographic form. This aimed at identifying the possibility of orthographic cues being embedded in the orthographic form of the target words (minimal pairs) in a way that L2 learners of English could retrieve this information for L2 vowel identification. It was hypothesised that the differences between the L1 and L2 orthographic systems could predict reliance on orthographic cues for the Greek L2 learners of English that could only been reflected if these possible cues were removed.

Experiment 3, a further exploration within the first strand, looked at the possibility of the consonantal environment of the target vowel offering potential cues that could be aiding vowel perception. This stemmed from previous research suggesting that, consonants surrounding a vowel could have an effect on the vowel quality (de Jong & Zawaydeh, 2002). Based on this assumption, it is hypothesized that identification and discrimination of vowels in the L2 English vowel contrast could be affected by consonant-related cues rather than vowel-related cues alone. To this end, Experiment 3 followed the same experimental design as in Experiment 1, but in this case the English vowel contrast involved controlled ambiguity of the consonantal environment.

The second strand of this thesis investigated brain plasticity and specifically age of acquisition-effects in relation with L2 experience. This was explored in Experiment 4 that tested learning outcomes following a perceptual training program based on high-variability phonetic training techniques. Child and adult L2 learners of English were tested before and after the training intervention in order to measure learning outcomes. The main research question here is to explore the link between the brain's early learning capacity and L2 experience, following perceptual training intervention. Success of previous training studies (e.g. Ylinen *et al*., 2010), allow the prediction that L2 learners will improve their perceptual identification and discrimination outcomes following perceptual training intervention. The fact that this study is the first to examine children as well as adults, aims to shed further light to L2 speech perception and implications on learning for adults and children alike.

One other area that is of particular interest in this case is whether children or adults will benefit more from perceptual training; i.e. whether adult L2 learners with much longer L2 experience could have an advantage over child L2 learners with little L2 experience. In other words, is it L2 experience that will lead adults to better learning outcomes? Could it be that children could demonstrate a more robust capacity for perceptually rearranging their phonetic cue weighting? Brain plasticity is explored in view of learning capabilities, L2 experience, age effects, and maturational constraints.

The third strand examines phonological processes in L2 reading and specifically looks into aspects of L2 orthography that could affect L2 phonological processing, especially in relation to spelling differences between the L1 and L2. Previous research has suggested an interaction between reading and phonological processing (e.g. Landerl *et al.*, 1997; Pattamadilok *et al.*, 2010; Adams, 1990; Badian, 1998; Bryant, 1986; de Jong & van der Leij, 2002; Denton *et al.*, 2000). Experiment 5 aimed at observing the use of orthographic cues which could be affecting L2 learners' phonological processing. This study is essentially a reading study that involved no auditory input. The aim was to observe the use of L1 orthographic cues at L2 reading and implications of reading processing for Greek learners of L2 English. Also, it examines whether Greek learners would demonstrate a tendency for orthographic transfer, given the orthographic differences between the L1 and L2 (discussed in section 1.6.3). It is proposed that reading outcomes can reveal phonological processes, given the orthographic differences between Greek and English. It is hypothesised that orthographic information included in the vowels of target English words can have valuable 'cues' for phonological processing that results in correct or incorrect reading performance for the Greek L2 learner of English. This study aims to explore the proposal that visual or orthographic cues may also be weighted by L2 learners differently to native speakers in the perception of speech sounds of a L2.

The predictions and hypotheses of the experiments described in the following chapters are summarised in Table 1.10 below.

Table 1.10: Predictions and hypotheses of the thesis.

Experiment	Research Question	Predictions/hypothesis
1 **(Chapter 2)**	1. Do Greek L2 learners of English utilise the same or different acoustic cues for perception of L2 sounds as English native speakers? 2. Do adult and child L2 learners of English perform differently in L2 speech perception due to L2 experience?	1. Greek L2 learners of English are expected to show perceptual assimilation and interference effects in relation to L2 cue-weighting. Native language magnet effects could also aid assimilation of two L2 categories into one native category. 2. Adult learners' experience with the L2 is expected to outperform child L2 learners.
2 **(Chapter 3)**	3. Do Greek learners of L2 English use orthographic cues for auditory identification?	3. It is hypothesized that orthographic cues embedded in words can aid auditory identification. Removal of such cues (e.g. use of pictures) could inhibit the possible use of orthographic cues available in words.
3 **(Chapter 3)**	4. Do consonants surrounding a vowel speech sound offer cues that could aid vowel identification or discrimination?	4. Consonantal environment could be affecting vowel perception, or provide cues for vowel identification or discrimination.
4 **(Chapter 4)**	5. Is there plasticity in perceptual learning for child and adult L2 learners alike? Are there child vs. adult age differences? 6. What are the benefits of the HVPT technique for L2 speech perception?	5. Plasticity in perceptual learning for child and adult L2 learners is predicted. 6. Previous research predicts that HVPT technique for L2 speech perception is beneficial for adult L2 learners. It is hypothesized that it will also be beneficial for child L2 learners.
5 **(Chapter 5)**	7. Are orthographic cues weighted by L2 learners differently to native speakers?	7. Greek learners of L2 English could demonstrate a tendency for orthographic transfer and use of L1 orthographic cues at L2 reading. It is predicted that orthographic cues are weighted differently by L2 learners to native speakers.

Chapter 2

Perceptual Identification and Discrimination of phonetic categories by Greek speakers of English.

The formation of native-like representations of second-language (L2) phonetic categories is a difficult task for the L2 learner, especially when phonetic cues are used differently between the native language (L1) and the L2 (e.g. Strange, Akahane-Yamada, Kubo, Trent, Nishi, & Jenkins, 1998; Flege, 1988; Flege, Bohn, & Jang, 1997; McAllister, Flege, & Piske, 2002; Ylinen *et al*., 2010; Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Bion, Escudero, Rauber, & Baptista, 2006; Rauber, Escudero, Bion, & Baptista, 2005; Escudero, Hayes-Harb, & Mitterer, 2008; Broersma, 2005; Cutler & Broersma, 2005). There are various kinds of cues that need to be weighted and integrated in order for a speech sound to be correctly identified (Holt & Lotto, 2006). Native speakers of a language learn to weight cues in way that those cues carrying more critical information are weighted higher than others (Holt & Lotto, 2006). A second-language (L2) learner's native language (L1), however, may inhibit correct cue weighting in the L2 by a process known as L1 transfer (Bohn, 1995; Strange, 1998). In his Perceptual Interference (PI) account (Iverson *et al*., 2003), Iverson argues that the acoustic cues that L2 learners use in practice are often different to those used by L1 speakers, thereby leading to inferior performance on perceptual categorization tasks. The cue weighting mismatch between an L1 and L2 could result in L2 learners identifying and relying on cues that are non-critical for identification or discrimination of phonetic segments in the L2 and treating such cues differently compared to a native speaker of the target L2. An example of this is given by Iverson *et al*. (2003), who studied the English /r/ and /l/ distinction by Japanese learners of

L2 English. In that study, it was shown that Japanese L2 learners of English relied on the F2 instead of the F3 formant values that were used as critical cues for the English L1 speakers.

Previous research (Best, 1995a; Best, McRoberts, & Goodell, 2001; Flege, 1995a, 2003) also suggests that the dissimilarity of cue weighting between an L1 and L2 could result in sounds being assimilated into the same L1 sound category. As Best's (1995a) Perceptual Assimilation Model (PAM; and PAM-L2, Best & Tyler, 2007) predicts, non-native sound categories typically assimilate into native categories and the discrimination of non-native phonetic contrasts can be accurately predicted by the perceived similarity between the contrasting sounds and the native categories.

Kuhl's (Kuhl, 1991; Kuhl *et al.*, 1992, 2008; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995) Native Language Magnet Model (NLM/NLM-e) proposes that future learning of L2 speech categories is constrained due to initial L1 sound mapping (i.e. committed neural structures) that has taken place. In this case, learning of new speech sounds of the L2 is predicted to be constrained as new sounds that are contained in the L2 input interfere with the already established mental maps of the L1 sounds. Stronger native language commitment is predicted for older rather than younger L2 learners. Although the NLM/NLM-e model proposes that the L2 learner can create and perceive L2 sound categories, the mechanism employed for the creation of L2 sound categories is not necessarily the same as the mechanism used for the creation and development of L1 sound categories.

Flege's (Flege, 1995a, 2003) Speech Learning Model (SLM) presumes that the learning capacity employed by infants and children when acquiring their L1 is retained by adult L2 learners who can achieve accurate perceptual learning of the L2 speech sound properties. Therefore the SLM predicts the formation of new L2 sound categories (e.g. Flege & MacKay, 2004). This L2 speech sound learning is based on the degree of perceived cross-language phonetic similarity: either due to the L2 learner's discrimination of the phonetic differences of two L2 sounds, or due to the existing phonetic differences between L1 and L2 sounds. It is predicted that new L2 sound categories are formed depending on the degree of phonetic dissimilarity between the L2 sound and its closest counterpart in the L1 (Flege, 2003). An example in support of this prediction is the claim that because the French /y/ sound is auditorily more distant from the French /u/ than the English /u/, native English speakers, learners of L2 French, were able to produce the French /y/ more accurately than the /u/ sound (Flege, 1987). In this manner, it is proposed that, although 'new' L2 sounds can be acquired, those sounds that are classified as 'similar' (between the L1 and the L2) are less likely to be acquired and developed to a native-like level. This is due to assimilation of the 'similar' L2 sound into an already established L1 sound category. A detailed explanation of each of the above models is provided in Chapter 1.

In this chapter[21], an experiment that explores perceptual identification and discrimination of English vowels by Greek native speakers is presented. Exactly which cues are critical for the speech sound identification and categorization in each language will vary. For example, the English tense-lax vowel contrast (/iː/ vs. /ɪ/) is a case where the vowel change incorporates both duration and spectral changes. In comparison, there is no such vowel distinction in Greek as the Greek five-vowel system is distinguished by place rather than manner of articulation, and there is only one /i/ speech sound category in the Greek vowel system.

As previously mentioned, one would expect that such a contrast might pose a problem for Greek speakers as in their L1 vowel duration is not used contrastively but rather as an effect of tempo or stress (Fourakis *et al*., 1999) compared with some other languages such as Finnish where duration is phonemically relevant in a way that vowel duration changes result in change of the word meaning (e.g. sika *vs.* siika) (c.f. Ylinen *et al.*, 2010). It could therefore be argued that Greek L2 speakers of English would identify contrastive phonological segments on the basis of the spectral qualities of the L2 vowels, whereas duration could be an irrelevant or misleading cue. On the

[21] Part of this chapter has appeared in Giannakopoulou, Uther, & Ylinen (2011).

other hand, it could be argued that L2 speakers may well attend to cues that are not existent in their L1.

Iverson *et al.* (2003) showed that Japanese speakers of L2 English attend to non-critical cues (i.e. F2 frequency instead of F3 frequency which is critical for English native speakers) when asked to distinguish the English /r/ - /l/ contrast. This is justified based on the fact that Japanese speakers use the F3 frequency for identifying their native /ɾ/ category which corresponds to the English /r/ - /l/ contrast. This supports McAllister *et al*. (2002)'s '*feature hypothesis'* which suggests that category formation is difficult when it makes use of a phonetic feature that is not used contrastively in the L1. McAllister *et al*. (2002)'s study compared perception of the Swedish vowel length contrasts by native speakers of American English, Latin American Spanish, and Estonian. Results showed that each language group performed based on whether durational features are used in their L1 or not. More specifically, Estonians that are most experienced with duration distinctions using them as a primary cue in their L1 performed better than the other two groups, the American English participants who are also experienced with duration distinctions using it as secondary cue in their L1 outperformed the Latin American Spanish group who have no duration distinctions for vowel perception in their L1. Therefore, relevant previous L1 experience with duration cues would determine the degree of successful learning of L2 vowels, also supported by the SLM (Flege, 1995a).

A contrastive view, however, comes from a study by Bohn (1995) that examined the perception of American English vowels marked by both duration and spectral differences, in German and Spanish speakers. German speakers were able to identify vowels by making use of spectral and duration cues that were present in German (Ramers, 1988) whereas L1 Spanish speakers relied on duration even though their L1 does not involve duration as a cue to indicate phonological contrasts (Navaro, 1968). Explaining these findings, Bohn (1995) proposed the 'desensitization hypothesis' which supports the view that L2 learners can be sensitive to duration cues (rather than spectral cues) even if duration is not used contrastively in their L1. Also, if L2 learners' previous linguistic experience did not sensitize them to spectral differences, then duration cues would be used for L2 vowel distinction. A number of subsequent studies have also supported this view that L2 learners tend to rely more on duration than spectral cues in the case of hard L2 contrasts (e.g. Escudero, 2005; Cebrian, 2006).

L2 learners use duration cues if spectral cues are unavailable, even when duration cues are not used contrastively in their L1. Lengeris (2009) examined how L1 duration experience can affect L2 vowel distinctions. He tested Greek and Japanese learners of L2 English. Even though duration is not used contrastively in Greek but it is used contrastively in Japanese, Lengeris (2009) found that both Greek and Japanese groups had access to duration cues at perceiving L2 English vowels. Assimilation patterns observed in his study suggest that duration interacted with spectral quality therefore the use of duration could not be attributed solely to salience. A further study by Lengeris & Hazan (2010) that performed phonetic training to Greek speakers of L2 English, suggests that L2 processing is affected by L1 experience but there are also individual differences observed as some individuals perform better at the use of L2 cues rather than being influenced by the L1.

Escudero & Boersma (2004), looking into the English /iː/ - /ɪ/ contrast distinction by Spanish speakers of L2 English, provided evidence of access to duration cues in English even if not relevant in Spanish. They suggested that for Spanish speakers it was easier to perceptually create a new vowel category (based on a single spectral category which is differentiated by duration, i.e. *long* /i/ and *short* /i/), rather than ‘splitting’ the already established Spanish /i/ based on spectral properties (i.e. /iː/ - /ɪ/) (Escudero & Boersma, 2004, p. 575). Support also comes from a study by Iverson & Evans (2007) who compared Spanish, French, German and Norwegian speakers’ vowel perception of English (Southern British English) vowels, and found that secondary acoustic cues (such as duration) were favored by L2 speakers irrespective of their L1. Similar effects of vowel duration on L2 perception have also been shown in a number of subsequent studies where duration is suggested to be a salient and easier cue to access. Therefore L1 experience with duration cues would not be obligatory for duration cue access in the L2 (e.g. Morrison, 2002; Escuderro &

Boersma, 2004; Escudero, 2005; Cebrian, 2006, 2007; Mora & Fullana, 2007; Cerviño & Mora, 2009). This also supports Bohn's (1995) idea that duration can be easily introduced as a cue in vowel perceptual contrasts even though it may not be present phonologically or contrastively in the learner's L1.

Therefore, Greek L2 learners of English may still access duration cues despite the fact that it would not result from L1 transfer. Given the various theories of speech acquisition, discussed in detail in Chapter 1, that emphasize the role of age of acquisition (Kuhl *et al*., 2008; Flege, 2002), performance of different L2 age groups were also explored in order to determine whether there were similar patterns of difficulty in adult versus child learners and examine maturational differences in acquisition of L2 phonetic segments.

2.1 Cue weighting between L1 Greek and L2 English

The case of Greek learners of L2 English represents a unique case that complements previous work on studying phonetic cue weighting in second language learners of English (e.g. Ylinen *et al*., 2010's study of Finnish speakers of English). As mentioned, Greek does not use duration to phonemically differentiate phonetic segments. It contains only one /iː/ category with spectral but not duration changes that correspond to the tense-lax English distinction /iː/ versus /ɪ/ with both spectral and duration (temporal) values characterizing each segment. In terms of categorization, and if critical cues are not weighted correctly by Greek learners of L2 English, it could be expected that the two L2 English vowel categories (i.e. /iː/ and /ɪ/) are assimilated into the one L1 /iː/ category as Best's (1995a) PAM (and PAM-L2, Best & Tyler, 2007) model mentioned above predicts. Performance in three types of tasks, perceptual identification, discrimination and classification was explored. The approach used by Ylinen *et al.* (2010) was adopted, i.e. stimuli were manipulated so that the use of the spectral cues is 'forced' in perception for L2 learners so that they were making their judgments on primary cues used by L1 speakers rather than on secondary cues. To this end, artificially modified stimuli were used so that the categorization of phonemic categories was not possible solely on the basis of the non-critical (duration) cues, but purely on spectral cues. Given that duration is not phonemically relevant within the Greek vowel inventory, the goal was to explore whether there was any effect of removing these cues in cue weighting and performance.

The first aim was to investigate the degree of perceptual identification and discrimination of the phonetic categories of the English tense-lax vowel contrast by Greek speakers of L2 English. The second aim was to explore whether the L1 affects the way in which acoustical cues (e.g. duration) are used in the L2. It also aims to explore relevant theories of cue-weighting in L2, e.g. feature hypothesis (McAllister *et al.*, 2002) and the desensitization hypothesis (Bohn, 1995) as discussed in Chapter 1. The third aim was to test whether age plays a significant role in the perception of L2 phonetic categories assessing performance of different age groups and compare L1 and L2 cue-weighting of English minimal pairs between Greek and English speakers. In order to assess cue weighting in L2 perceptual identification and discrimination tasks and the role of different cues (e.g. spectral and duration) in the English tense-lax distinction in perceptual processing, responses to stimuli with normal versus modified durations were compared. It is hypothesized here that Greek L2 speakers of English could demonstrate an ability to identify contrastive phonological segments by processing spectral cues of the L2 vowels with duration being perceived as an irrelevant or misleading cue or may as well attend to cues that are not existent in their L1 (as proposed by Bohn, 1995) but are not critical. In the case of English native speakers, it is expected that spectral cues are weighted as critical (Flege & Hillenbrand, 1986; Holt & Lotto, 2006).

2.2 Experiment 1: L2 perceptual identification and discrimination

2.2.1 Method

The purpose of this experiment was to test the perceptual identification and discrimination of the tense-lax vowel distinction by Greek speakers of L2 English. In order to assess accuracy levels, categorical identification and discrimination tasks were designed accordingly following categorical perception studies that have previously used this paradigm (e.g. Levy & Strange, 2008; Gottfried, 1984). Cue weighting was tested and three different tasks were designed in order that different aspects of it may be examined: one task where auditory stimuli were presented in isolation (i.e. minimal pair counterparts) and only visually presented as minimal pairs (perceptual identification), one where the comparison was made only on the basis of auditory percepts (auditory discrimination) and another where single counterparts of a minimal pair were to be classified in a relevant category (auditory classification). In addition to assessing age effects or maturational constraints, cue-weighting between Greek and English speakers was explored, and the additional aim was to test whether the existence of context (i.e. visual in the case of perceptual identification or auditory by listening to both minimal pair counterparts and discriminating between the two) or the lack of it (i.e. classifying single words) could affect participants' performance as well as explore on perceptual identification and discrimination patterns between the two language groups (Greek and English).

2.2.1.1 Participants

Adult groups**:** Twenty adult native speakers of Standard Modern Greek (8 female, 12 male) aged 20-30 (mean age = 25.4) were recruited from the University of Patras, Greece and twenty monolingual English native speakers (14 female, 6 male) aged 19-26 (mean age = 21.4) were recruited from Brunel University, West London, UK. The English native speakers served as controls. The Greek speakers had all lived in Greece and had all studied English as L2 in school (public and private education; L2 English education mean = 8.7 years). Their level of proficiency in English was advanced in listening, speaking, reading and writing (all had received the Cambridge Certificate in Advanced English or Certificate of Proficiency in English) as recorded in the language background questionnaire that all participants completed before testing. None had spent more than 2 weeks in an English-speaking environment. Greek learners of L2 English received a small gift of stationery items as reward for their participation and were entered into a prize draw. English native speakers received course credit for their participation. All adult participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

Child groups**:** Forty-five children were tested, all native speakers of Standard Modern Greek. They were arranged in three age groups: fifteen aged 7-8 years (8 female, 7 male; mean age = 7.6; L2 English education mean = 1.2 years); fifteen aged 9-10 years (7 female and 8 male; mean age = 9.8; L2 English education mean = 3.7 years); fifteen aged 11-13 years (9 female, 6 male; mean age = 12.3; L2 English education mean = 5.4). They had all lived in Greece and were students of L2 English (public and private education). Their respective level of proficiency in English was basic, lower intermediate and intermediate in listening, speaking, reading and writing as recorded in the language background questionnaire that all participants completed before testing. None had spent more than 2 weeks in an English-speaking environment. They were recruited from a private language school in Greece. Forty-five child monolingual native speakers of Standard English were also tested and served as controls: fifteen aged 7-8 (6 female, 9 male; mean age = 7.4); fifteen aged 9-

10 (10 female and 5 male; mean age = 9.5); fifteen aged 11-13 years (6 female and 9 male; mean age = 11.9). They were recruited from a local primary school. All child participants received stationery as reward for their participation. All child participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

2.2.1.2 Stimuli

The auditory stimuli[22] were words pronounced by a male native speaker of English representing typical Southern British English pronunciation (37 years of age). He was asked to read a set list of words a number of times (the first and last words in the list were excluded in order to avoid certain list-reading properties) and the utterances were recorded digitally in an acoustically attenuated room. One exemplar for each word sample (the best exemplar representing typical English pronunciation) was selected.

Forty-five minimal pairs (i.e. word pairs that are differentiated by one single vowel sound, e.g. *sit* versus *seat*) with the English tense-lax vowel distinction were used (see Appendix 1). All auditory stimuli were natural speech items with normal vowel durations.

A second set of minimal pairs was created where the normal vowel duration was digitally manipulated with Praat software (Boersma and Weenink, 2004) so that for each minimal pair the intrinsically long /iː/ vowel was shortened to match the duration of the intrinsically shorter /ɪ/ vowel and vice versa (see Figure 2.1).

[22] The same stimuli have been previously used in Ylinen *et al.* (2010).

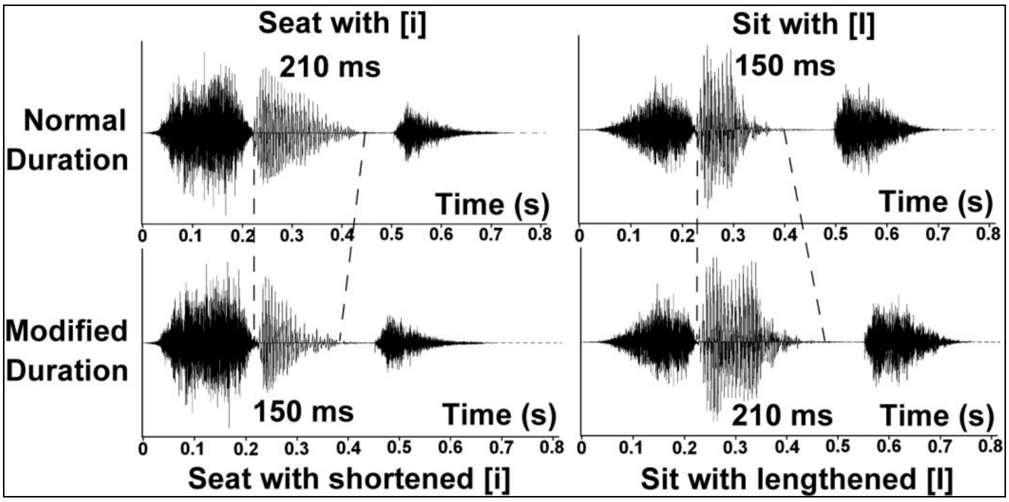


Figure 2.1: An example of the stimuli used. Top: The minimal pair ‘seat’ and ‘sit’ with normal duration (i.e., as ordinarily pronounced by a native English speaker). Bottom: The same minimal pair with modified vowel durations: [iː] was shortened to correspond to the original duration of [ɪ], and [ɪ] was lengthened to correspond to the original duration of [iː].

For example, for the minimal pair *sit* – *seat*, the sound stimulus would be either [siːt] as in *seat* but the long vowel /iː/ was shortened to match the duration of the short vowel /ɪ/ or [sɪːt] as in *sit* but the short vowel [ɪ] was lengthened to match the duration of the long vowel /iː/. Vowel modifications were resynthesized with the Pitch-Synchronous Overlap and Add (PSOLA) technique (Boersma and Weenink, 2004). In these modified stimuli, the spectral quality of vowels was preserved even though duration cues were ambiguous (thus not available) (see Figure 2.1) which forced participants attend to spectral cues only. This was one of the main aims of these experiments, discussed further in later sections.

Stimuli were high-pass filtered at 80 Hz in order to eliminate any low-frequency noise, the stimuli intensity was normalized (in this case the peak was scaled to 90% maximum), and the initial and final 2ms of each recorded stimulus were ramped in order to eliminate any click artifacts. The fundamental frequency (F0) of the stimuli used varied within natural range while the maximal values at the beginning of the vowel was 121-201 Hz (mean 151 Hz) and the minimal values at the end was 102-168 Hz (mean 121 Hz).

Thus, 90 digitally recorded words[23] arranged as 45 minimal pairs were used as auditory stimuli throughout this experiment and the respective 45 minimal word pairs (orthographic form) were used as visual stimuli. The auditory manipulation that resulted in the modified duration words produced a second set of 45 auditory minimal pairs while visual stimuli were not altered or modified in any way.

2.2.1.3 Apparatus

Stimuli were presented on a laptop (AMD Sempron) with E-Prime software (Schneider, Eschman, & Zuccolotto, 2002a,b). All auditory stimuli were binaurally presented through high quality headphones (SONY MDR-V150) at 44 kHz, 16-bit resolution and at a comfortable listening level (varying between 65-75 dB). Visual stimuli were displayed on a 33 x 20 cm monitor. Reaction times (RT) and responses were automatically recorded for each participant through the E-Run[24] software application. Participants responded by pressing an allowable relevant button on the computer keyboard which triggered the next trial after a 1000 ms delay.

2.2.1.4 Procedure

This experiment included two tasks with two conditions. The tasks and stimuli within each task were presented in random order. Participants were tested individually. They completed a language background questionnaire, gave

---

[23] All auditory and visual stimuli used in this experiment have also been previously used in experimental tasks described in Ylinen et al. (2010).
[24] The E-Run software application is the presentation component of E-Prime 1.1 enabling for millisecond precision of stimulus presentation, as well as synchronizations of stimuli (e.g. visual and/or auditory) and data collection.

written consent and participated in all tasks. Each session lasted approximately 45 minutes.

Perceptual Identification Task (PI): This task used 45 normal vowel duration and 45 modified vowel duration minimal pairs of English words arranged into two conditions (natural and modified duration stimuli, as described in Section 2.1). Each trial consisted of an auditory stimulus (i.e. the auditory form of one of the minimal pair counterparts) presented through the headphones and a simultaneous visual stimulus (i.e. the orthographic form of the minimal word pair, e.g. *sit* – *seat*) presented on the screen (see Figure 2.2). In both conditions, each auditory word was presented once in random order, accompanied by the visual stimulus consisting of the orthographic representation of the corresponding minimal pair on the screen. For example, the minimal pair *sit* – *seat* appeared twice visually but the first time the participants heard '*sit*' and in a later presentation, they heard its minimal pair counterpart, '*seat*'. Thus, 90 stimuli were presented for each condition, 180 in total (45 minimal pairs x 2 tokens =90 stimuli x 2 conditions = 180 stimuli). For each participant, all stimuli were presented in random order arranged automatically by E-Prime (Schneider, *et al.*, 2002a,b). The two counterparts (visual stimuli) of each minimal pair were also presented in random order. Participants were instructed to choose which one of the two words presented on the screen they heard through pressing a relevant key on the computer keyboard. Due to the nature of the stimuli in the modified duration stimuli condition, participants could only base their choice on spectral rather than durational cues to make the vowel identification.

Figure 2.2: Screen shot of a random perceptual identification (PI) task example. The sound effect feature here is only to indicate the auditory stimulus accompanying the on-screen visual stimulus however it was only presented aurally and did not appear on the participants' screen.

***Auditory Discrimination Task (AB-X)*:** The stimuli were presented only aurally as a discrimination task, arranged in an ABX format (i.e. ABA or ABB). Each trial comprised a sequence of a minimal word pair (word A followed by word B or word B followed by word A) and a third stimulus being the exact repetition of either word A or word B (see Figure 2.3). Two conditions were included: a natural stimuli condition and a modified duration condition (with stimuli matched in duration to the other word in the minimal pair) and each condition consisted of 45 trials. Thus, in one condition (natural stimuli type) natural vowel length stimuli were used, in another condition (modified duration stimuli type) both natural and modified duration vowels were used so that in each auditory sequence the vowel duration for either word A or B was a modified vowel duration stimulus type that matched in duration its naturally long or short vowel counterpart. For example, [siːt] – [sɪːt] – [siːt] or [sɪːt] – [siːt] – [sɪːt] or [sɪt] – [sit] – [sɪt] or [sɪt] – [sit] – [sit]. In this case, the spectral qualities were preserved although the durations of the vowels across all three words of the sequence in each trial were identical. Participants were instructed to respond by pressing relevant keys on the computer keyboard whether the third word on each trial was same as the first word or same as the second word in the auditory sequence.

Figure 2.3: Screen shot of a random auditory discrimination (AB-X) task example. The sound effect features here indicate the three auditory stimuli which were presented aurally and did not appear on the participants' screen.

***Classification Task*:** This task involved natural and modified duration (i.e. shortened or lengthened vowel duration type) stimuli. Each trial consisted of a single auditory word stimulus, of either normal or modified duration vowel type. Ninety stimuli were presented (3 types x 30 pairs of each type = 90 stimuli). This was again an auditory task. The screen displayed three categories: 'normal', 'shorter' or 'longer' throughout the task. Participants were instructed to classify the vowel of each auditory word stimulus into one of the three categories displayed on the screen by pressing a relevant key on the computer keyboard that represented each category. They selected 'normal' if the vowel sounded normal or natural, 'shorter' if it sounded shorter-than-normal, and 'longer' if it sounded 'longer-than-normal' (see Figure 2.4).

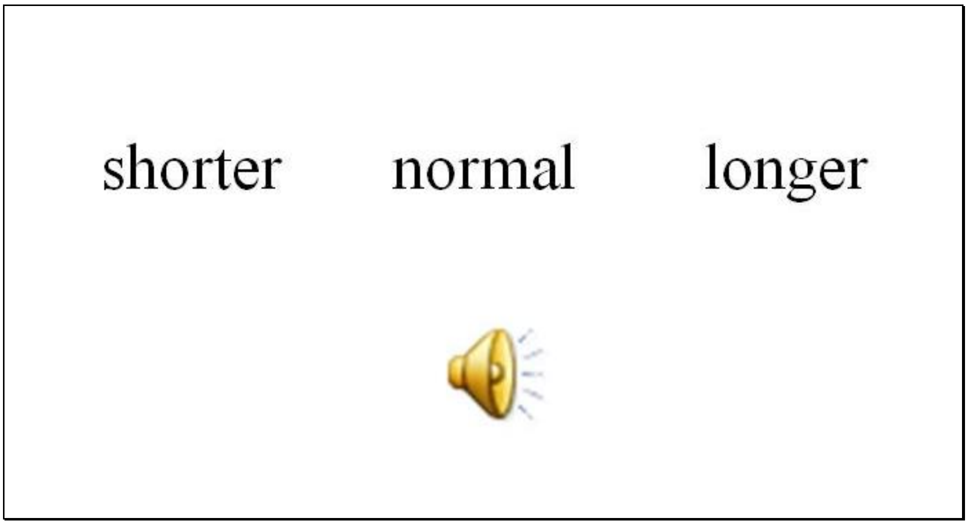


Figure 2.4: Screen shot of a random auditory Classification task example. The three words ‘shorter – normal – longer’ were the visual stimulus across all task trials and served as the three categories in which sounds were to be classified. The sound effect feature here indicates that one auditory stimulus was presented aurally in each trial and did not appear on the participants’ screen.

2.2.2 Results

For the perceptual identification tasks, correct responses were averaged across participants for the two language groups, age groups and task conditions. To examine accuracy performance of perceptually identifying the correct vowel sound, correct responses were analysed with a three-way mixed design analysis of variance (ANOVA) with Language Group and Age as between subject factors and Vowel Length (duration) as within subject factors. Greeks were less accurate than English speakers in the perceptual identification task and the modified duration condition in particular (see Figure 2.5 and Table 2.1). The natural stimuli task showed an effect of age development, with Greek adults performing better than Greek children and English adults showing ceiling effects, outperforming all other groups.

English speakers performed better overall than Greek speakers with a significant main effect of Language group ($F_{(1,\ 122)}$= 1244.469, p< .001), see Figure 2.5 and Table 2.1. Greek child participant scores were at chance level for all perceptual identification task conditions. There was a significant main effect of age ($F_{(3, 122)}$= 16.659, p< .001, see Figure 2.5 and Table 2.1) with Bonferroni contrasts showing that adult participants performed better compared with child participants for the respective language groups, but with no statistically significant difference between younger and older children. There was a significant effect of Vowel Length ($F_{(1, 122)}$= 118.862, p<.001), see Figure 2.5 and Table 2.1, showing that identification of words with natural vowel length was better than those with modified duration vowels and confirming the assumption that vowel duration was indeed used as a critical cue for perceptual identification performance.

There was also a significant Age x Language interaction ($F_{(3, 122)}$= 4.347, p<.05), see Figure 2.5 and Table 2.1, showing that there was less difference between child and adult participants in the English group compared to a more marked difference between child and adult participants in the Greek group, where Greek children performed at chance level. The Vowel Length x Age interaction was significant ($F_{(3, 122)}$= 9.610, p< .001), see Figure 2.5 and Table 2.1, because the adult groups performed considerably better than all child groups in the natural duration condition, whereas in the modified duration condition the adult groups' performance was comparable to the child groups. Specifically, the decrease in performance between natural and modified duration was greater for the adults. Concordantly, the three way interaction was significant ($F_{(3, 122)}$= 12.537, p< .001).

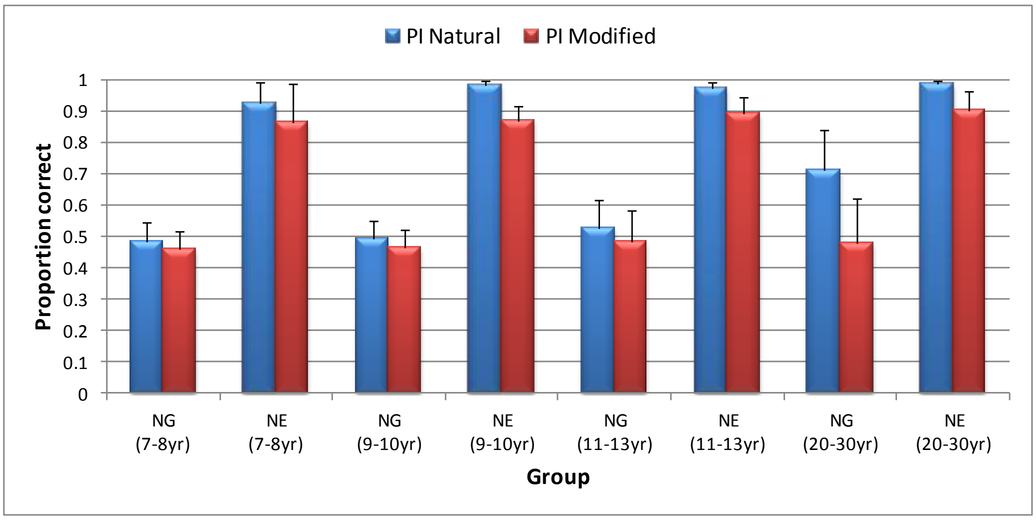


Figure 2.5: Accuracy levels for native Greek (NG) and native English (NE) child and adult participant groups for the Perceptual Identification (PI) task for natural and modified conditions.

Table 2.1: Mean accuracy scores and SD for adult and child groups for Perceptual Identification (PI) for Natural (Nat) and Modified (Mod) duration condition.

Group	PI Nat n=45		PI Mod n=45	
	Mean	SD	Mean	SD
Greek child, 7-8 (n=15)	0.49	0.06	0.46	0.05
English child, 7-8 (n=15)	0.93	0.07	0.86	0.12
Greek child, 9-10 (n=15)	0.49	0.06	0.47	0.05
English child, 9-10 (n=15)	0.98	0.04	0.87	0.05
Greek child, 11-13 (n=15)	0.52	0.09	0.48	0.10
English child, 11-13 (n=15)	0.97	0.02	0.89	0.05
Greek adult, 20-30 (n=20)	0.71	0.13	0.48	0.14
English adult, 20-30 (n=20)	0.99	0.02	0.90	0.06
Mean	0.76	0.06	0.68	0.08

For the AB-X discrimination task, correct responses were averaged across participants for the two language groups, age groups and task conditions. Ceiling effects were observed for most groups for the AB-X discrimination tasks (natural and modified duration conditions), with English participants outperforming the Greeks. Greek adult participants also show higher scores for the AB-X task compared with the ABX modified duration task where duration cues had been equalized (see section 2.2.1). Responses were analysed with a three-way mixed design of variance (ANOVA) with Language Group and Age as between subject factors and Vowel Length (duration) as within subject factors.

There was a main effect of language ($F_{(1, 122)}$= 406.707, $p < .001$), see Figure 2.6 and Table 2.2, where English performed better overall compared with Greek speakers. For both language groups, children performed worse than adults ($F_{(3, 122)}$= 3.876, $p<.05$), see Figure 2.6 and Table 2.2. For all age groups, discrimination of words with natural vowel duration was overall better compared with words with modified vowel duration as revealed by a main effect of Vowel Length ($F_{(1, 122)}$= 30.609, $p<.001$), see Figure 2.6 and Table 2.2. A significant Language x Age interaction ($F_{(3, 122)}$= 8.871, $p< .001$, see Figure 2.6 and Table 2.2) showed that English adults performed better than Greek adults and English child participants performed better than Greek child participants. A Language x Vowel Length interaction revealed that Greeks performed significantly better with natural vowel duration words ($F_{(1, 122)}$= 32.554, $p<.001$, see Figure 2.6 and Table 2.2) confirming the assumption that Greek participants weight duration cues as more primary compared with spectral cues. The 3-way interaction was significant ($F_{(3, 122)}$=4.561, $p< .05$, see Figure 2.6 and Table 2.2) since all groups performed better with the natural vowel duration words compared with modified duration words. However, the largest difference was revealed with Greek adults whose performance with words of modified vowel duration was impaired.

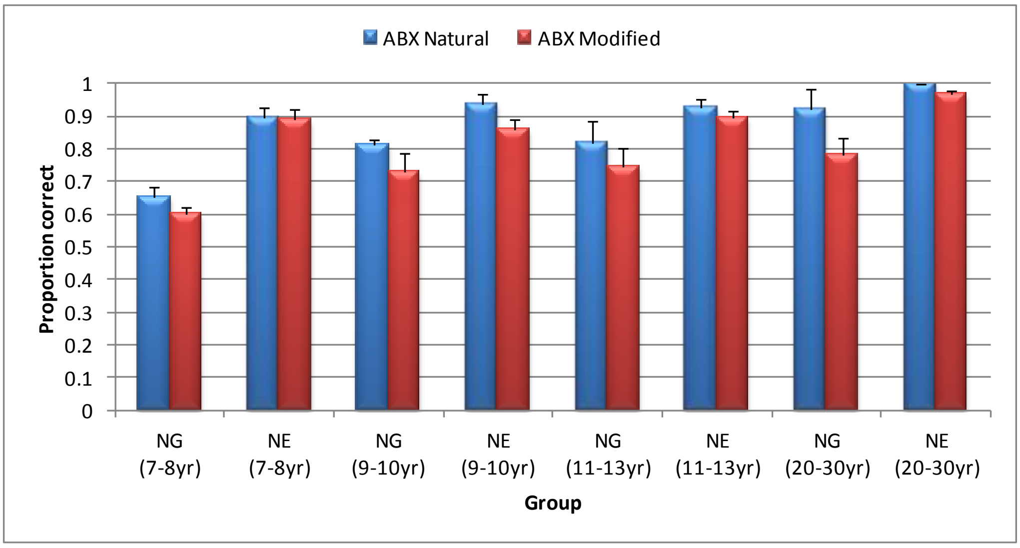


Figure 2.6: Accuracy levels for native Greek (NG) and native English (NE) child and adult participant groups for the ABX task (natural and modified conditions).

Table 2.2: Mean accuracy scores and SD for adult and child groups for Auditory Discrimination (AB-X) for Natural (Nat) and Modified (Mod) duration condition.

Group	ABX Nat n=45		ABX Mod n=45	
	Mean	SD	Mean	SD
Greek child, 7-8 (n=15)	0.651	0.03	0.74	0.02
English child, 7-8 (n=15)	0.907	0.03	0.892	0.03
Greek child, 9-10 (n=15)	0.814	0.02	0.762	0.06
English child, 9-10 (n=15)	0.896	0.03	0.858	0.03
Greek child, 11-13 (n=15)	0.798	0.07	0.814	0.06
English child, 11-13 (n=15)	0.925	0.03	0.896	0.02
Greek adult, 20-30 (n=20)	0.92	0.07	0.73	0.05
English adult, 20-30 (n=20)	0.997	0.06	0.967	0.06
Mean	0.86	0.04	0.83	0.04

For the classification task, although all English groups performed better compared with Greek groups, performance was in general low (at chance level) across all age groups. This may be affected by the task requirements and is discussed in section 2.5. A two-way univariate analysis of variance (ANOVA) with two independent variables (age and language) was conducted for the perceptual classification task. The analysis yielded a significant main effect of language ($F_{(1,\ 127)}$=85.114, p< .001, see Figure 2.7) showing that English participants performed better than Greek participants overall but no effect of age was observed. No other main effects or interactions were significant.

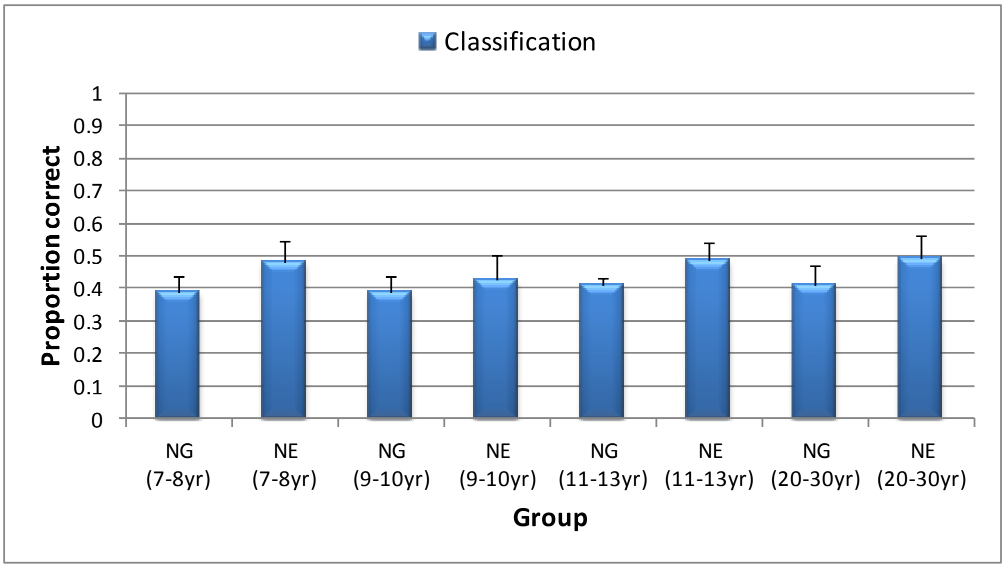


Figure 2.7: Accuracy levels for native Greek (NG) and native English (NE) child and adult participant groups for the Classification task.

2.3 General Discussion

This experiment explored the perceptual identification and discrimination of L2 vowels with cues that were not phonemically relevant in the L1. Specifically, Greek speakers of L2 English were tested with respect to cue weighting for the English /iː/ and /ɪ/ (tense-lax) distinction. Three experimental task types were designed featuring two task conditions. The same participants were tested in all tasks. The goal of this experiment was firstly to assess the degree of perceptual identification and discrimination of phonetic categories; secondly to examine the use of spectral and duration cues and how L1 may affect L2 cue weighting for the Greek L2 learners of English; thirdly, to observe any age-related effects and how different age groups would perceptually respond to certain L2 phonetic categories.

Access to durational cues by speakers with no such L1 experience has been a much discussed issue in recent years and various hypotheses have been proposed. As mentioned earlier, two theories in this field are the *feature hypothesis* (McAllister *et al*., 2002) which proposes that L2 learners do not have access to L2 cues that are not used contrastively in their L1 and that therefore L1 transfer is possible, and also the *desensitization hypothesis* (Bohn, 1995) which proposes that durational cues can be accessed by L2 learners even if not used contrastively in their L1 (in particular for the case of L2 vowels) and duration being an easier cue to access compared to spectral cues (Ylinen, 2005; Escudero, 2005; Cebrian, 2006, and others). In order to test the assumptions of these hypotheses, Experiment 1 presented in this chapter used stimuli with natural and modified duration vowels in order to control for the cues being used by participants; i.e., Greek L2 learners of English and English native controls alike.

It is clear that L2 cues are weighted differently between Greek L2 speakers of English and native English speakers as the experimental results reveal, in support of the Perceptual Interference account (Iverson *et al.*, 2003). L2 learners show preference in certain phonetic cues (duration vs. spectral) confirming perceptual interference between L1 and L2 phoneme categories. This limited access to all L2 cues complements the assumption that certain L2 categories assimilate into L1 categories, as proposed by Best's PAM/PAM-L2 (Best, 1995a; Best & Tyler, 2007). Experiment 1 demonstrated that although native English speakers use spectral features as the primary cue and duration as secondary cue, Greek speakers of L2 English use duration as the primary cue in the recognition of English vowels while performance is impaired for those tasks where identification and discrimination is based solely on spectral cues.

These data appear to favour the desensitization hypothesis and data reported by Bohn (1995) over the feature hypothesis proposal (McAllister *et al*., 2002).

Even though a duration cue may not be phonemically relevant in the L1, it appears to be perceptually identified and treated as a primary cue in the identification and discrimination of L2 vowel quality. This holds true for both Greek adult and child groups although Greek adults show a greater tendency to rely on duration cues. The spectral vowel qualities that are present do not seem perceptually detected with identification task results at chance level, compared to ceiling effects by the native English groups. This indicates that cues are weighted differently between the L1 and L2 English groups, but also native language cues (e.g. primary cues in Greek as L1) do not always predict cue weighting in the L2. There may be several possible reasons for this. It could be a consequence of the duration changes that are naturally found within Greek vowels, even though they are not phonemically relevant. A study by Fourakis *et al.* (1999) reveals that the same Greek vowels can be measured as having longer or shorter durations depending on whether they are produced in a stressed *vs*. unstressed, focus *vs*. nonfocus, or slow *vs*. fast manner. They found very significant duration differences between unstressed and stressed nonfocus vowels with the latter being the longest. This suggests that some sort of duration change is naturally found in Greek vowels (being affected by type of speech, e.g. giving a lecture or speaking to a friend were the speech types that Fourakis *et al.* (1999) tested), therefore even though not phonemically relevant, duration changes are in fact used to determine cues for other features such as stress.

Another explanation could be that duration is simply an easier discrimination to make than a new kind of spectral change. Or, it could be a combination of both duration and spectral change, an explanation that is also supported by Lengeris (2009). Lengeris (2009) tested Greek learners of L2 English. However, he examined a larger number of vowel contrasts (compared to the study presented in this chapter that examined one vowel contrast, i.e. tense – lax), using temporal manipulations that aimed to explore the use and reliance on duration as a cue. Lengeris (2009) found that 'English vowels generally fitted better to L1 categories in the context where they resembled more the duration of the spectrally closest L1 vowel' (p. 183). These findings challenge

Bohn's (1995) assumptions of the desensitisation hypothesis and therefore the use of duration simply as a more salient cue. However, the use of the modified duration stimuli, particularly for the perceptual identification task in Experiment 1 and the results that emerged for the L2 speakers of English compared with the native English speakers, reveal that there is a greater reliance on duration cues by L2 speakers than the Lengeris (2009) study suggests. Given the experimental differences, however, between the Lengeris (2009) study and the current experiment, it could be suggested that further research would have to be conducted to determine which of these explanations are the most likely. Nonetheless, the main conclusion that L2 speakers of English are dependent on different cues to that of L1 speakers of English is still supported.

Four age groups were considered in order to identify any age-related effects. The perceptual identification (PI) task reveals a difference between the L2 child groups (where all perform at chance level) and the L2 adult group. In comparison, a slightly different pattern is observed for all four age groups for the auditory discrimination (AB-X) task, with adults performing better than all three child groups and older child participants performing better compared with younger ones. Differences in L2 proficiency, however, should also be considered in relation to age-related effects. The fact that L2 groups had different L2 proficiency starting points could be argued that it could interact with any underpinning age-effects. Therefore, even though there is a clear pattern of results observed where performance is better by older than younger participants, further observations would be necessary in order to control for the additional factor of proficiency levels (i.e. L2 experience). For example, it could be interesting in future experiments to test child and adult participants who had the same amount of exposure to the L2.

Another observation could be in relation to task-type and age-related performance: despite the L2 child groups' performance at chance level in the case of the perceptual identification task, the higher scores obtained in the discrimination task show that task-type could also interact with age as well as proficiency. These data could also be interpreted in terms of speech learning models that suggest a consolidation of neural and perceptual phonetic categories with increasing age (Kuhl, 2008; Flege, 2002). Considering this assumption along with cue weighting results discussed above, it could be suggested that durational cues are perceptually identified as primary cues, L2 learners tend to utilize durational cues and these are established even more with increasing age.

Also, even though all Greek adult and child participants rely primarily on duration cues (as in the modified vowel duration tasks all groups perform worse compared with natural vowel duration), Greek adults seem to rely more heavily on duration cues and this is true for all experimental tasks (vowel identification and discrimination tasks alike). Thus, it could be suggested that weighting duration as a critical cue appears to be perceptually developed by L2 speakers over time and the more experience and exposure to the L2 they have the more duration cues are weighted as primary. Such explanation is particularly justified by L2 adult performance in the perceptual identification task conditions (natural versus modified) where in the case of modified stimuli condition L2 adults perform at chance levels compared with higher scores for the natural stimuli condition.

An additional issue that should be discussed here is potential effects of lexical bias or knowledge on identification results, especially for children. Lexical bias has been known in the literature as a factor affecting phonological categorisation in spoken word recognition tasks. Initial observations by Ganong (1980) showed that listeners shifted their phonetic category boundary locations for /t/ versus /d/ based on the lexical status of the stimulus items. For example, listeners selected /t/ more frequently than /d/ for the *task – dask* continuum, and selected /d/ more frequently than /t/ for the *tash – dash* continuum. This effect reflects an advantage for real words over nonwords in the case of lexical identification tasks with ambiguities in the phonetic category level. Such lexical bias effect could be interpreted as the result of the robustness of the human speech recognition system to acoustic variability. That is, ambiguous acoustic input could result in the speech recognition system to be biased toward matching this input to an existing rather than non-existing entry (i.e. a word vs. nonword) in the mental lexicon.

Given this observation, that has since been supported by a variety of similar studies (e.g. Fox, 1984; Pitt, 1995; Walley & Flege, 1999; Thompson & Hazan, 2010), such lexical bias effect could also be applied in a cross-cultural context. Lexical knowledge of the L2 word stimuli used in the experiment presented in this chapter, could be suggested to have a potential effect in word identification. The lexical bias effect observed by Ganong (1980) in the case of words over nonwords, could be applied in the case of L2 known words versus L2 unknown words. For example, if one of the two words of a minimal pair happens to be an unknown word to the L2 learner, in view of the 'Ganong effect' considered here, the L2 learner could erroneously select the known over the unknown word presented to them (equivalent to word vs. nonword bias effects). In the experiment presented in this chapter, Greek L2 speakers' scores for the perceptual identification task being mostly at chance level could support this suggestive potential effect. However, the auditory discrimination task opposes or eliminates the possibility of lexical bias on the basis of higher scores obtained by the same L2 participants. Of course, as already discussed, task type could also play an additional role.

It could be suggested that this is a limitation for this study and word knowledge could have been tested for L1 and L2 speakers alike prior to their participation in the experimental tasks presented in this chapter. This way it would have been made possible to take into account or exclude the possibility of any lexical bias based on lexical knowledge. However, this is an issue that Experiment 2 (Chapter 3) is taking into consideration and lexical knowledge effects are discussed in sections 3.1.3 and 3.1.4.

One further issue that needs to be considered in relation to the experiment presented in this chapter is how the orthographic representation (visual stimulus) of the English word stimuli could perceptually affect auditory identification. For example, for the natural duration stimuli task in the case of the 'seat-sit' visual stimulus, when the auditory stimulus is 'seat' Greek L2

English speakers may be choosing ‘seat’ (which would be the correct answer) based on the orthography of the word (being an orthographically ‘longer’ word) compared to the case of the auditory stimulus being ‘sit’ and L2 English speakers choosing ‘sit’ for their response (which would be the correct response), again, based on the orthographic representation of the word ‘sit’ simply being an orthographically ‘shorter’ word. If this pattern is followed, then results should be distorted for the modified duration condition, while ceiling effects would be expected to be observed for the natural duration condition. However, this was not the case as results for the natural duration condition were also impaired for Greek L2 learners of English. These implications are addressed in Chapter 3 and 5 with a variety of tasks aiming to shed more light on this very issue.

Chapter 3

Perceptual Identification and Discrimination of phonetic categories using Picture and Cross Spliced stimuli pairs.

The two experiments presented in this chapter are designed to complement the data from Experiment 1 described in Chapter 2. Perceptual identification and discrimination of phonetic categories is again explored, but with a view to re-designing the tasks in order to control for two areas that were not possible to explore through the previous experimental tasks: 1. the orthographic representation of word stimuli could have affected (to some degree) participants' choice for the perceptual identification task, particularly for non-native language learners and, 2. the consonant 'frame' (C_C) surrounding each vowel under consideration could provide essential information toward cue weighting features of the vowel itself. These two issues are addressed and explored separately in two experiments described below.

3.1 Experiment 2: Picture Stimuli

3.1.1 Introduction

Orthographic representation is examined because in any of the minimal pairs used in Experiment 1 (Chapter 2), the word that contains the long vowel /iː/ is usually spelt with a greater number of letters than its counterpart that contains the short vowel /ɪ/. Such examples are: *seat* vs. *sit, bean* vs. *bin, feet* vs. *fit, sheep* vs. *ship,* and more. As discussed in Chapter 2, Greek participants (both child and adult groups) seem to rely on duration as a primary cue for perceptual identification tasks. Duration is known to have a quantitative feature due to its temporal dimension. This is why vowels are often categorized based on their temporal 'length' and are referred to as long or short vowels (e.g. Klatt, 1976; Hillenbrand, Getty, Clark & Wheeler, 1995; Hillenbrand, Clark & Nearey, 2001). This may give rise to the idea that participants may assume that more letters means longer vowel duration. Therefore, if Greek speakers tend to rely on duration for vowel categorization (long vs. short), then the number of letters in the word stimulus they are presented with could be an additional 'cue' for this categorization. In such cases, when presented with a word stimulus *seat – sit* on the screen and hear [siːt] as the auditory stimulus for this pair, they could be selecting the word 'seat' merely based on the fact that the long-sounding vowel should be matched with the word containing 'more' letters, and in the case they hear [sɪt] then the word choice would be 'sit' based on the fact that the short-sounding vowel should be matched with the word containing 'less' letters. The main aim of this experiment is therefore to explore the degree of interference of orthographic cues with perceptual identification task performance.

Even though some word minimal pairs included the same number of letters for each counterpart (e.g. beater – bitter), the vowel itself was still spelled with either one or two graphemes. That is, for the minimal pair counterparts that involved a 'long' vowel sound, there were three different spellings for the /i:/ sound: 'ee', 'ea', or 'ei'. For the 'short' vowel minimal pair counterparts, there was one spelling for the /ɪ/ sound: 'i' (see Appendix 2a). This is what is referred here as 'more' versus 'less' letters contained in the minimal pair words. Nevertheless, in order to also control for any spelling effects of the words' general spelling as opposed to spelling of the vowel itself which is the main target for the purposes of this experiment, data analysis will also include tests that will aim to show whether performance was higher for the minimal pairs that included orthographic length cues (e.g. seat- sit) than for those that did not (e.g. beater-bitter).

In order to control for the possibility that orthographic representation of the word stimuli provides cues for perceptual identification and thus affects and distorts the experimental results, this experiment explores perceptual identification with the use of pictures as visual stimuli as opposed to word stimuli. Therefore, instead of using minimal pairs in their orthographic form (word stimuli), the same minimal word pairs were replaced by relevant pictures that resembled the meaning of the respective words. For example, in the case of the minimal pair 'sheep – *ship*', the two pictures were used as visual stimuli instead of words in their orthographic form, as in the example below (see Figure 3.1; a full list of stimuli is also given in Appendix 2a):

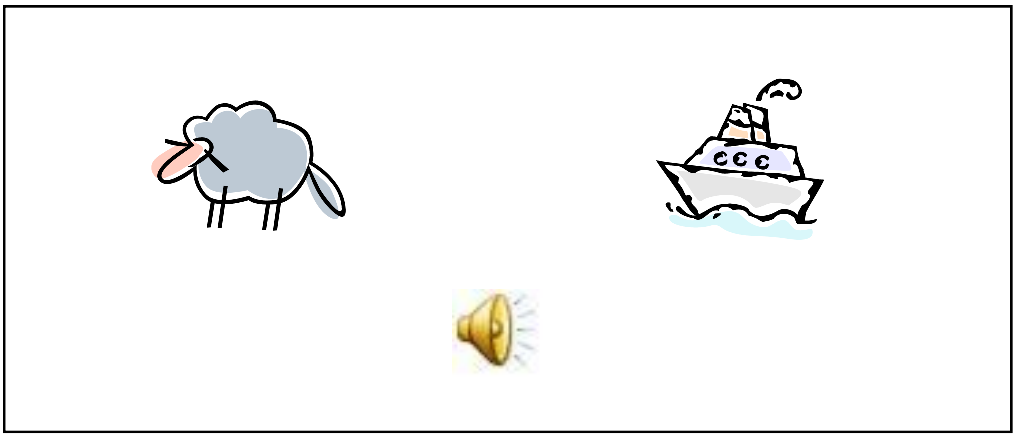


Figure 3.1: Screen shot of a random picture perceptual identification (PI) task example representing the 'sheep – ship' word stimulus pair. The sound effect feature here is only to indicate the auditory stimulus accompanying the on-screen visual stimulus however it was only presented aurally and did not appear visually on the participants' screen.

The auditory stimuli set used throughout the experimental tasks, was identical to that used in Experiment 1 (Chapter 2). This experimental paradigm allowed for perceptual identification of the auditory stimulus excluding any potential orthographic cues since the pictures used did not contain orthographic information.

Picture naming is of course essentially a translation task, and therefore involves different processes to those described in Experiment 1 (Chapter 2). In a study by Potter, So, Von Eckhardt, & Feldman (1984), there is evidence that translation resembles picture naming and it is proposed that translation is conceptually mediated. Picture naming involves retrieving the picture's concept. Hence translation inevitably involves concept retrieval as part of the translation process. Indeed, the direction of the translation (whether from L1 to L2) may also affect the degree of concept mediation. Kroll & Stewart (1994) used a bilingual translation experiment which included naming and translation experimental conditions. For the naming condition task, Dutch speakers of L2 English were instructed to read out loud the word that would appear on the computer screen, while the voice output was recorded by a voice-operated relay. The stimuli were English and Dutch words and appeared in random order. For the translation condition task, the same participants followed the same procedure but instead of reading the onscreen stimulus word, they were instructed to translate it into the opposite language (i.e. Dutch or English). Kroll & Stewart (1994) found that L1 to L2 translation was conceptually mediated whereas L2 to L1 translation was reliably faster and more accurate leading to the proposal that L2 to L1 translation is lexically mediated (i.e. based on connections between words in the two languages). This suggests a different translation route, depending on the direction (L1 to L2 or, L2 to L1) of translation, may be employed.

This study, although it does not aim to test the above assumptions directly, aims at considering perceptual identification processes in the case of L2 picture naming and L2 word translation. Specifically it is hypothesized that: if the auditory stimulus has been conceptually established in the case of Greek L2 English learners, then it may be possible to correctly identify and choose the relevant picture (from the represented pictorial minimal pair). If the previous orthographic representation of the minimal word pair provided some additional cues, then performance in this task could be impaired. Also, in case the picture stimuli were difficult or non prototypical representatives for the underlying word meanings, results would also reveal low scores. Low scores would also be expected in case either age group (child or adult group) did not know the L1 meaning of the L2 words when asked to translate the L2 words under consideration into equivalent L1 words.

Considering age group differences, there could be an expectation that children perform worse in general than adults in a reading task due to a slower reading ability (e.g. Carter, Mintun, & Cohen, 1995; Balaban, Snidman, & Kagan, 1997; Cornelissen, Munro, Fouler, & Stein, 1993; and for numerical processing see, Girelli, Lucangeli, & Butterworth, 2000), although they might perform better with a picture task given that reading skills are not directly necessary for this task type. On the other hand, given that this is essentially a translation task, it is likely that there could be an effect on children's performance (i.e. giving lower scores) if they have to mentally transform a visual representation to text.

The use of pictures as visual stimuli in this experiment, thus, required control for the following: 1. Participants' familiarity with the concepts underlying the pictures used; 2. Participants' familiarity with the connection of L2 words and their L1 equivalents (for all the meanings represented by the pictures). This would indicate the degree of participants' access to the semantic inventory as well as access to the conceptual inventory for adult and child groups alike. In order to control for the above, this experiment involved three stages: a. prototypicality test for the pictures used (this relates to the selection of the picture stimuli, discussed in section 3.1.2.1); b. paper-based test with two tasks, a picture match task and a translation task (experimental task 1); c. the computer-based perceptual identification task with natural and modified duration conditions (experimental task 2). More detail on the experimental procedure is provided in the Method section 3.1.2).

The short paper-based tasks served as control tasks prior to the computer-based tasks and further aimed to: 1. help participants familiarize themselves with the picture stimuli used in the computer-based tasks; 2. ensure participants' ability to correctly identify which picture refers to which word (through the picture match task); and, 3. measure the degree to which participants are familiar with the word stimuli used (through the translation task) which would also control for the degree of any lexical bias/knowledge (e.g. Ganong, 1980; Fox, 1984; Pitt, 1995; Walley & Flege, 1999; Thompson & Hazan, 2010). The computer-based tasks formed the main experimental tasks and aimed to: 1. explore the degree of reliance on duration cues, 2. the degree of reliance on orthographic cues. The main research question of this experiment is whether Greek learners of L2 English use orthographic cues for auditory perceptual identification, in relation to duration cues (available or non-available) in the auditory stimulus.

It is also important to mention here the impact of task difficulty in comparison with the original task as presented in the previous chapter (in the case of the orthographic visual stimuli). Learners of L2 English are usually instructed using the orthographic forms of English words in relation to relevant L1 words. Therefore, pictures are not necessarily registered as word representations for L2 words. For example, the (verb) word ‘sit’ is usually learnt as an action or command word compared to the noun ‘seat’ which may have an object as mental representation. It is therefore essential to control for the extent to which participants can accurately relate the picture stimuli to the respective words, despite their grammatical status (e.g. verb, noun, proper noun, etc). The lack of orthographic representations (or orthographic cues) as well as ‘forced’ access to conceptual representations, which resembles a translation task, could be regarded as contributing factors to the task difficulty.

3.1.2 Method

3.1.2.1 Picture Stimuli

For the list of minimal word pairs used in Chapter 2, a list of respective picture pairs was created. The pictures were clip art files found on the internet via the Microsoft Clip Art Web application. Due to the nature of some of the original list of words (i.e. the meaning of some words being abstract, e.g. ‘sin’) it was not possible to represent all minimal word pairs in picture form. This resulted in a set of 33 picture pairs (i.e. 66 picture stimuli).

In order to confirm the prototypicality[25] of the pictures representing the respective words they were presented in the form of a questionnaire to 10

[25] For a critical discussion on prototypicality with respect to word meaning see Giannakopoulou (2003). For further discussion also see: Rosch, 1978; Barsalou, 1983; Geeraerts, 1986; Laurence & Margolis, 1999; Lehrer, 1989.

English and 10 Greek native speakers (age 20-30 years[26]) who were asked to judge how good representative each picture was for the respective word using a five-level Likert scale (1=poor, 5=very good). Pictures that did not receive scores above 3 were excluded from the picture stimuli set used in the picture experiment. This aimed to eliminate the use of pictures that could be non-prototypical or not representative enough for the purposes of this experiment. This resulted in the reduction of the initial number of picture stimuli to 30 picture pairs being used (i.e. 60 picture stimuli). Appendix 2a presents the picture stimuli used for the purpose of this experiment.

Age of acquisition (AoA) and frequency (written) scores for the 60 words that were included in the experimental tasks were examined. Scores were obtained from the MRC database (Wilson, 1988) and Bristol Norms (Stadthagen-Gonzalez & Davis, 2006) (see Table 3.1). AoA and frequency scores were not available for all words used in the experimental tasks. For the words where there were available AoA scores, it was revealed that words were normally acquired by native English speakers before the age of 7 (therefore these words could be considered as acquired by child participants in these experimental tasks since they were 7-8 years of age).

Even though the limited database information available for the list of words used throughout this experiment poses a potential methodological weakness, the experimental tasks did not exclude any words merely based on lack of AoA or frequency scores. The main rationale for this was that such exclusion would result in a limited number of word /picture stimuli. Also, results from a smaller dataset would limit the discussion in relation to other studies described in this thesis. However, a repeat analysis excluding the items that do not have AoA and/or frequency data is reported (see section 3.1.3).

[26] Participants who took part in the visual stimuli selection stage did not participate in any further experimental tasks.

Table 3.1: Age of acquisition (AoA) and frequency (KFFR, i.e. Kucera and Francis written frequency) scores obtained from MRC Database (Wilson, 1988) and Bristol Norms (Stadthagen-Gonzalez & Davis, 2006; scores shown in red). Words included in later AoA analyses are marked with asterisk.

WORD	AOA	KFFRQ	WORD	AOA	KFFRQ
BEAN	-	5	BIN*	1.9	9
BEAT	-	68	BIT	-	101
BEAD	-	1	BID*	4.2	425
BEAKER*	5.2	2	BICKER	-	-
BEATER	-	-	BITTER*	6.7	53
BEEF*	5.1	32	BIFF	-	-
BEES	-	-	BIZ	-	2
CHEAP*	5.6	24	CHIP*	4.6	17
DEEP	-	109	DIP	3.3	6
FEET	-	283	FIT	-	75
FEAST*	3.5	3	FIST*	5.1	26
GENE	-	9	GIN	-	23
HEAT*	2.8	97	HIT	-	115
LEAD*	4	129	LID	-	19
LEAFED	-	1	LIFT*	3.1	23
LEAK*	3.4	2	LICK	-	3
LITRE	-	-	LITTER*	5.4	3
NEAT*	5.4	21	KNIT	-	10
PEACH*	2.9	3	PITCH	-	22
PEAK*	6.8	16	PICK*	4.8	55
PETE	-	-	PIT*	3.1	14
SEAT*	4.5	54	SIT	-	67
SEED*	5	41	SID	-	-
SEEK*	5.8	69	SICK*	4	51
SHEEP*	3.9	23	SHIP*	2.4	83
SLEEP	-	65	SLIP	-	19
SLEET	-	1	SLIT*	3.6	6
SPEAK	-	110	SPICK	-	-
TEAM*	5.4	83	TIM	-	-
TEEN	-	6	TIN*	2.8	12

3.1.2.2 Participants

Adult groups: Ten native adult speakers of Standard Modern Greek (6 female, 4 male) aged 19-30 (mean age = 24.3) were tested. They had all lived in Greece and had all studied English as L2 in school (public and private education; L2 English education ranged 8-9 years; mean = 8.5 years). Their level of proficiency in English was advanced in listening, speaking, reading and writing (all had received the Cambridge Certificate in Advanced English or Certificate of Proficiency in English) as recorded in the language background questionnaire that all participants completed before testing. None had spent more than 2 weeks in an English-speaking environment. Ten monolingual English native speakers (5 female, 5 male) aged 19-30 (mean age = 21.4) were also tested and served as controls. All adult participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

Child groups: Ten child native speakers of Standard Modern Greek (5 female, 5 male), aged 7-8 (mean age = 7.8) were tested. They were recruited from a private language school in Greece. They had all lived in Greece and were students of L2 English (public and private education). Their respective level of proficiency in English was basic in listening, speaking, reading and writing as recorded in the language background questionnaire that all participants completed before testing (L2 English education ranged 1-2 years; mean = 1.6 years). None had spent more than 2 weeks in an English-speaking environment. Ten child monolingual native speakers of Standard English were also tested and served as controls (6 female, 4 male; mean age = 7.9). They were recruited from local primary schools in the UK. All child participants received stationery as reward for their participation. All child participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

3.1.2.3 Apparatus

Stimuli were presented on a laptop (AMD Sempron) with E-Prime software (Schneider *et al*. 2002a,b). All auditory stimuli were binaurally presented through high quality headphones (SONY MDR-V150) at 44 kHz, 16-bit resolution and at a comfortable listening level (varying between 65-75 dB). Visual stimuli were displayed on a 33 x 20 cm monitor. Reaction times (RT) and responses were automatically recorded for each participant through the E-Run[27] software application. Participants responded by pressing a button on the computer keyboard which triggered the next trial after a 1000 ms delay.

3.1.2.4 Procedure

This experiment included two main tasks: a paper-based task, and a computer-based perceptual identification task. Greek adult and child participants took part in all tasks (paper-based and computer-based tasks). English adult and child participants took part in the computer-based tasks only. Computer-based tasks were counterbalanced. Participants were tested individually. Each session lasted approximately 30 minutes for the adult groups and approximately 45 minutes for the child groups. Prior to the tasks, all participants (parents or carers in the case of child participants) signed a consent form and completed a language background questionnaire. Ethical Approval had also been obtained prior to conducting the study.

Prior to the paper-based task, participants were allowed familiarization time with the picture materials. Each participant was provided with the list of pictures and the respective words these pictures resembled (e.g. the word '*ship*' paired with the respective picture resembling a *ship*). The material (pictures and words) contained in this list was the same as the experimental material used in the paper-based and computer-based tasks. The picture-word pairs were printed in random order and not arranged as minimal pairs. Participants were

[27] The E-Run software application is the presentation component of E-Prime 1.1 enabling for millisecond precision of stimulus presentation, as well as synchronizations of stimuli (e.g. visual and/or auditory) and data collection.

instructed to spend approximately 10-15 minutes checking the list of picture-word pairs, and familiarize themselves with the material. This procedure aimed at eliminating any erroneous performance in the paper-based and computer-based experimental tasks due to previous lack of exposure to the conceptual connections of the specific word - picture pairs used throughout this experiment.

Picture Naming Task and Translation Task (control, paper-based task): Prior to the computer-based tasks, Greek participants (child and adult groups) took part in two short paper-based tasks: they were provided with a list of printed pictures (the same as those used in the experimental tasks involving picture stimuli) arranged randomly, and a numbered list of the respective English words also presented in random order on a separate piece of paper. They were asked: 1. to match the English words to the pictures by noting the word number next to the matching picture; 2. to translate the list of English words into Greek by providing a Greek definition for each English word.

Perceptual Identification Task (PI) with picture stimuli: The PI task procedure followed is the same as described in Chapter 2, however fewer auditory stimuli were used since some of the abstract words used in the previous experiment could not be pictorially represented easily. This task used 30 natural vowel duration and 30 modified vowel duration minimal pairs of English words arranged into two conditions (natural and modified duration stimuli respectively) and the respective picture stimuli (e.g. for the auditory stimulus '*seat*', one picture representing 'seat' and one picture representing 'sit' was displaying on the computer screen as visual stimuli and the same visual stimuli would randomly appear throughout the task in the case of the auditory stimulus being '*sit*'). Thus, 60 auditory word stimuli were presented for each condition, 120 in total (30 minimal pairs x 2 tokens =60 stimuli x 2 conditions = 120 stimuli). For each participant, all stimuli were presented in random order arranged automatically by E-Prime (Schneider, *et al*., 2002a,b). The two pictures (visual stimuli) of each minimal pair were also presented in random order (on the left or right side of the screen). Participants were instructed to choose which one of the two pictures presented on the screen they heard through pressing a relevant key on the computer keyboard. Due to the nature of the stimuli in the modified duration stimuli condition, participants could only base their choice on spectral rather than durational cues to make the vowel identification.

Perceptual Identification (PI) Task with word stimuli (control task): A PI task was also administered, an exact repetition of the PI task described in Chapter 2 (again, using natural and modified duration auditory stimuli arranged in two conditions, natural and modified duration), 30 minimal pairs were presented in each condition. However, no picture stimuli were used in this case. The aim of this task was to serve as a control task for comparison reasons since the participants in the study of Chapter 2 were different from those described here.

3.1.3 Results

Figure 3.2 shows the accuracy scores for Greek adult and Greek child groups for the Picture Match task and the Translation task. Adult participants did slightly better than child participants in both tasks (See Appendix 2b for individual scores of adult and child participants).

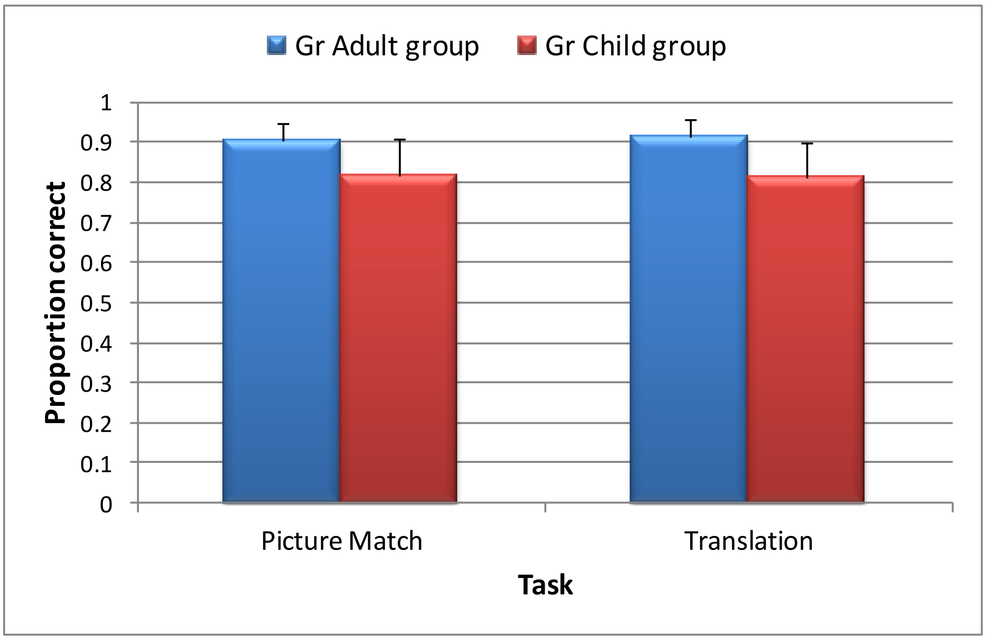


Figure 3.2: Accuracy scores for Greek adult and Greek child groups for the Picture Match task and Translation task.

Results were analysed with a one-way ANOVA with Age group (2 levels) as factor. There was a significant main effect of age for both tasks, Picture Match task ($F_{(1, 18)}$ = 6.918, $p < .05$) and Translation task ($F_{(1, 18)}$ = 10.918, $p< .05$). Despite the significant effect of age, both groups still achieved high scores (over 80%) which suggests that both groups were aware of the meaning of the words used in the experimental sets and could match the words with the pictures with a high level of accuracy.

Accuracy scores for all computer-based tasks were analysed with a three-way mixed design ANOVA with Language Group and Age as between subject factors and Vowel Length (duration) as a within subject factor. English child and adult participants did better than Greek child and adult participants. Task condition also affects performance with the natural duration condition acquiring higher accuracy scores than the modified duration condition. Figure 3.3 shows the accuracy scores for all groups and conditions.

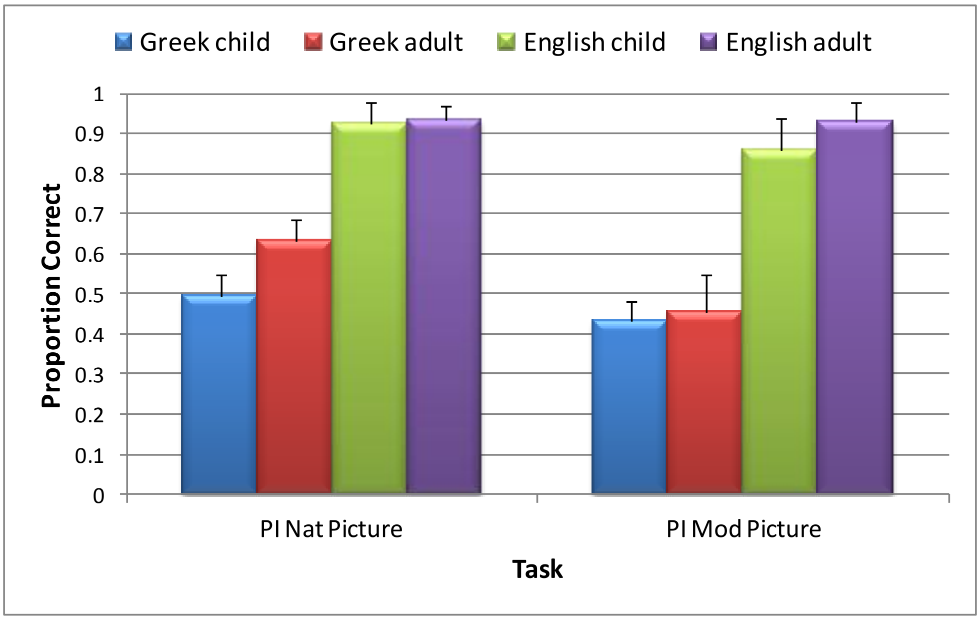


Figure 3.3: Accuracy scores for adult and child groups for the Perceptual Identification (PI) Picture task for Natural (Nat) and Modified (Mod) duration condition.

There was a significant main effect of Language ($F_{(1, 36)}$ = 619.895, p< .001), see Figure 3.3, where English groups performed significantly better than Greek groups. There was a main effect of Age ($F_{(1, 36)}$ = 17.430, p< .001), see Figure 3.3, showing that adults performed better than child participants. There was a main effect of Vowel-length ($F_{(1, 36)}$ = 56.542, p< .001), see Figure 3.3, because performance was better in the natural duration condition than in the modified duration condition. A Vowel-length x Language interaction ($F_{(1, 36)}$ = 15.050, p< .001, see Figure 3.3) emerged because Greek participants performed better in the natural duration condition than the modified duration condition, whereas English participants showed less marked difference in performance between duration conditions. This again confirms that duration was used as a primary cue by Greek participants unlike English participants. No other interaction effects were significant.

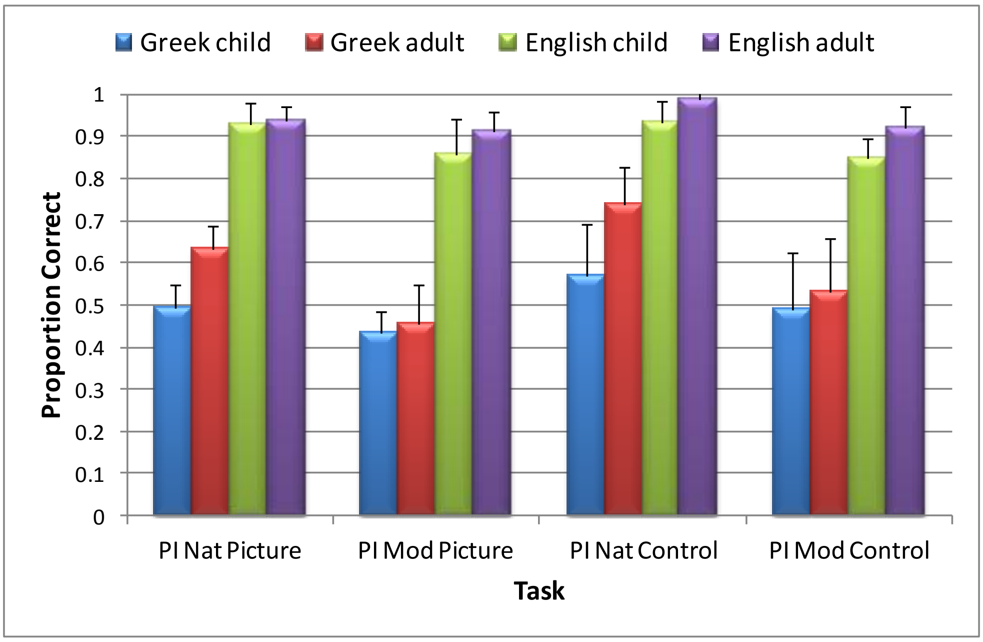


Figure 3.4: Accuracy scores for adult and child groups for the Perceptual Identification (PI) Picture and Control task for Natural (Nat) and Modified (Mod) duration condition.

Table 3.2: Mean accuracy scores and SD for adult and child groups for Perceptual Identification (PI) Picture and control tasks for Natural (Nat) and Modified (Mod) duration condition.

Group	PI Nat Picture n=30		PI Mod Picture n=30		PI Nat Control n=30		PI Mod Control n=30	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek child (n=10)	0.49	0.06	0.43	0.05	0.57	0.13	0.49	0.14
Greek adult (n=10)	0.63	0.05	0.46	0.09	0.74	0.09	0.53	0.13
English child (n=10)	0.93	0.05	0.86	0.08	0.93	0.07	0.85	0.05
English adult (n=10)	0.94	0.03	0.93	0.05	0.99	0.01	0.91	0.05
Mean	0.75	0.05	0.67	0.07	0.81	0.07	0.69	0.09

Figure 3.4 and Table 3.2 show both PI task types (picture stimuli and control (word) stimuli task) for all participants and task conditions. English groups performed better than Greek groups across all tasks. The main effect of Language ($F_{(1, 36)} = 479.613$, $p< .001$), see Figure 3.4 and Table 3.2, was significant, showing that English speakers performed better than Greek

speakers overall. The main effect of task (PI Picture task vs. PI Control task) was significant and Greek groups performed worse in the Picture compared to the control PI task ($F_{(1, 36)} = 11.110$, $p < .05$, see Figure 3.4 and Table 3.2). This suggests that the idea that the word orthography aids perceptual identification working as an additional 'cue', may be supported. The same analysis was repeated excluding any items with no AoA or low frequency scores (see Table 3.1) and this showed a similar pattern of results. An additional analysis was also repeated in order to see whether performance was higher for the minimal pairs that included orthographic length cues (e.g. seat – sit) than for those that did not (e.g. beater – bitter), but there were no significant results between the two minimal pair types for the Picture stimuli task (p = .321) and the Control task (p = .599) alike. There were no other significant interactions.

3.1.4 Discussion

This experiment explored the degree of reliance on duration cues and orthographic cues. This was tested through the use of perceptual identification tasks, using natural and modified duration auditory stimuli, and the replacement of the previously used orthographic representations of word minimal pairs as visual stimuli with picture pairs. Prototypicality of the pictures used in the experimental tasks was measured using data from a small experiment testing native English and Greek learners of L2 English, which resulted in the exclusion of those pictures that did not receive high prototypicality scores (see section 3.1.2.1). Also, in order to control for participants' familiarity with the concepts underlying the pictures used and participants' familiarity with the connection of L2 words and their L1 equivalents (for all the meanings represented by the pictures), time for familiarization with the experimental material was allowed prior to the start of the experimental tasks and two paper-based tasks were then administered: a picture-match task and a translation task. Only Greek learners of L2 English participated in these paper-based tasks as it was necessary to assess the degree of access to semantic and conceptual inventories for Greek adult and child groups in terms of recognizing the relevant meaning associated with pictures presented to them and how these relate to L2 words. The paper-based tasks also served as a screening test toward any potential effects of lexical bias for the L2 learners on the basis of known versus unknown words. This has already been discussed in Chapter 2 (Experiment 1) that identified this issue as a limitation and this is one of the reasons that Experiment 2 presented in this chapter aimed to control for.

Results revealed high scores on performance with the picture-match task and translation task, with the Greek adult group performing slightly better than the Greek child group. High scores for both tasks indicate that Greek adult and child participants were aware of the meaning associated with the picture stimuli used and were also able to translate the L2 words into their L1 with a high degree of accuracy (raw data presented in Appendix 2b). This was done to ensure familiarity with the concepts and the word meanings used throughout this experiment as well as controlling for any potential effects of lexical bias that would be revealed if there was a large proportion of unknown words used as experimental stimuli. However, no effects of lexical bias or knowledge were observed. Kroll & Stewart (1994)'s proposal that L2 to L1 translation is lexically mediated is also supported by the high scores in the translation task where L2 words were translated into L1 equivalents. The fact that translation has been suggested to resemble picture naming (Potter *et al.*, 1984), and that it is a conceptually mediated task, again high scores in the picture-match task reveal that the concepts or meanings resembled by the picture stimuli had been mentally established for both Greek adult and child groups.

The computer-based perceptual identification task results reveal that adult groups do better than child groups within their respective language groups; however English child groups outperform both Greek adult and child groups. It is, however, necessary to point out that Greek child participant scores were largely settled at chance levels. Despite the fact that Greek children's performance in perceptual identification tasks with natural vowel duration stimuli was generally better than modified vowel duration stimuli (which suggests reliance to duration cues), their generally low scores also suggest that there may be implications of age as well as proficiency, since child and adult participants differed at both age and proficiency level. Therefore, this combination of factors should be taken into account for a more accurate result interpretation.

Also, Chapter 2 (Experiment 1) discussed differences in performance between perceptual identification and auditory discrimination task, particularly for child participants. What was suggested in Experiment 1 was that task type could also influence scores and general performance for children since children performed better at auditory discrimination task than perceptual identification task. An auditory discrimination task could not be practically administered as part of this experiment since the main aim was to test the use of visual cues (i.e. pictures versus words). Task difficulty should nevertheless be taken into consideration, particularly for L2 child groups.

The relatively high scores achieved by child participants in the paper-based task for picture match and translation, point to the exclusion of lexical bias or knowledge effects and suggest that child participants were familiar with the picture – word minimal pairs used (which posed a limitation for Experiment 1 due to the fact that lexical knowledge effects had not been controlled for or tested in that experimental paradigm). However, the low scores (at chance level) by Greek child participants in the perceptual identification tasks in the current experiment (particularly for the picture task) suggest an interference between reliance to orthographic cues and also possible effects of age versus proficiency levels in comparison to Greek adult participants (specifically for the picture and word tasks with natural vowel duration stimuli).

The fact that the Greek adult group performs better in the PI control task which involved words as visual stimuli compared to the PI task which used picture stimuli (although auditory stimuli used were the same for both tasks), suggests that the orthographic representation of words could be providing additional cues that were not available in the case of picture stimuli. Ceiling effects observed for English adult and child groups across tasks and conditions reveal that English native speakers do not use orthographic cues in order to accurately identify auditory stimuli (despite the fact that in the case of word stimuli participants have to make a choice between two orthographic representations). Additionally, results confirm outcomes of Experiment 1 (Chapter 2) that duration is not used by English native speakers as a primary cue for correct perceptual identification, unlike Greek learners of L2 English who show preference to duration cues rather than spectral cues for correct vowel identification.

3.2 Experiment 3: Cross Spliced Stimuli

3.2.1 Introduction

Another possible confounding factor that needed to be explored with this experiment was the possibility that the consonants surrounding the vowel in each word stimulus would affect perceptual identification or discrimination as opposed to a perceptual choice being made upon the vowel alone. In other words, the consonants may be carrying important information or 'cues' that could be affecting perceptual identification of the vowel constituent. Such an assumption could challenge vowel cue weighting claims and assumptions. It has been suggested that consonants surrounding a vowel could generally affect the vowel quality (de Jong & Zawaydeh, 2002). The easiest way to test whether this would have had an impact on the results in Experiment 1 (Chapter 2) is to use stimuli which had auditory stimuli pairs cross spliced (the technique is described in detail in section 3.2.2 below). It should be noted here that this chapter does not aim to address the general issue of how consonants affect vowels based on the consonant voicing, place and manner of articulation as well as co-articulation features in a CVC environment (e.g. Jacewicz & Fox, 2008; de Jong, 2004; de Jong & Zawaydeh, 2002). It only explores the case of specific minimal word pairs used for the purpose of the current study and how manipulating the C_C environment of each vowel could affect vowel perception. The research question specifically explored here is whether consonants surrounding a vowel speech sound offer cues that could aid vowel identification or discrimination. The general aim of this experiment is to investigate whether results concur with the perceptual identification and discrimination patterns as discussed in Chapter 2 or the use of different stimuli type (e.g. cross spliced pairs) could affect participants' performance and accuracy levels and therefore lead to different conclusions.

3.2.2 Method

3.2.2.1 Cross Spliced Stimuli

The auditory word stimuli used in the current experiment were pronounced by a male native speaker of English representing typical Southern British English pronunciation and were identical to those used in Chapter 2 (i.e. forty-five minimal pairs with the English tense-lax vowel distinction, see Appendix 1). Both sets of natural and modified duration vowel pairs were used.

The pairs were then digitally manipulated (using the Adobe Audition 3.0 software application) so as to produce a set of cross spliced pairs: the vowel of each minimal pair counterpart was selected, removed and pasted between the consonant space of the other minimal pair counterpart and vice versa. The end-result was a new set of minimal pairs with the vowels having being swapped while the neighbouring consonants remained intact (see Figure 3.5). This procedure used both natural duration and modified duration stimuli, the same stimuli used in Experiment 1 and presented in Chapter 2, thus one new set was created with cross-spliced natural duration stimuli and one new set with cross-spliced modified duration stimuli.

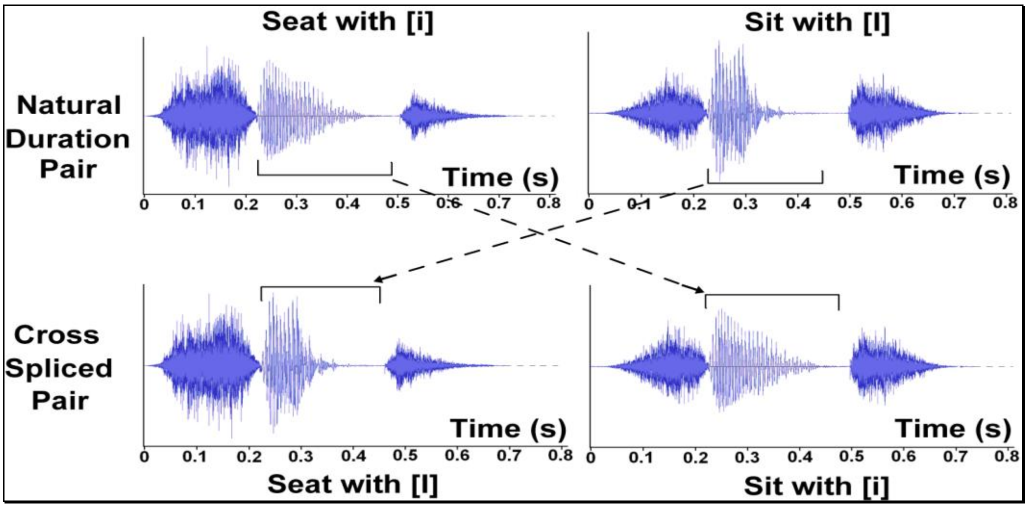


Figure 3.5: An example of the cross spliced stimuli. Top: The minimal pair ‘seat’ and ‘sit’ with natural duration (i.e., as ordinarily pronounced by a native English speaker). Bottom: The same minimal pair with cross spliced vowels: [iː] was replaced by [ɪ] for ‘seat’ and [ɪ] was replaced by [iː] for ‘sit’ while surrounding consonants remained the same respectively as shown in the above figure.

It is possible that the cross splicing manipulation procedure can result in stimuli that sound unnatural as reported by Kemps, Ernestus, Schreuder, & Baayen (2005). In order to avoid any such effects, the stimuli were tested by ten English native speakers who were unaware of the auditory manipulation. The native English speakers confirmed that all stimuli sounded native-like, albeit with a somewhat synthesized quality. It should be emphasized that the splicing was done so as to minimize the effect of the co-articulation to the listener’s ear, and at the same time maintain as natural stimuli as possible.

3.2.2.2 Participants

All participants were the same as in the Picture Stimuli Experiment, but this study was conducted in a different testing session.

3.2.2.3 Apparatus

Stimuli were presented on a laptop (AMD Sempron) with E-Prime software (Schneider *et al*. 2002a,b). All auditory stimuli were binaurally presented through high quality headphones (SONY MDR-V150) at 44 kHz, 16-bit resolution and at a comfortable listening level (varying between 65-75 dB). Visual stimuli were displayed on a 33 x 20 cm monitor. Reaction times (RT) and responses were automatically recorded for each participant through the E-Run software application. Participants responded by pressing an allowable relevant button on the computer keyboard which triggered the next trial after a 1000 ms delay.

3.2.2.4 Procedure

This experiment included two tasks: a perceptual identification (PI) task and an auditory discrimination (AB-X) task. An additional control task was also administered. Participants were tested individually. Each session lasted approximately 30 minutes. Prior to the computer-based task, all participants (parents/main caregivers in the case of child participants) signed a consent form and completed a language background questionnaire.

Perceptual Identification (PI) Task with Cross Spliced word stimuli: The PI task procedure and participant instruction followed is the same as described in previous PI tasks in Experiment 1 (Chapter 2), using words as visual stimuli and cross spliced words as auditory stimuli. This task used 45 natural vowel duration and 45 modified vowel duration minimal pairs of English words arranged into two conditions (natural and modified duration stimuli respectively). Thus, 90 auditory word stimuli were presented for each condition, 180 in total (45 minimal pairs x 2 tokens = 90 stimuli x 2 conditions = 180 stimuli). For each participant, all stimuli were presented in random order arranged automatically by E-Prime (Schneider *et al.*, 2002a,b). Participants were instructed to choose which one of the two words presented on the screen they heard through pressing a relevant key on the computer keyboard. Due to the nature of the stimuli in the modified duration stimuli condition, participants could only base their choice on spectral rather than durational cues to make the vowel identification.

***Auditory Discrimination (AB-X) Task with Cross Spliced word stimuli*:** The stimuli were presented only aurally as a discrimination task, arranged in an AB-X format (i.e. ABA or ABB). Each trial comprised a sequence of a minimal word pair (word A followed by word B or word B followed by word A) and a third auditory word stimulus being the exact repetition of either word A or word B. This task followed the same design as the AB-X task in Chapter 2, however stimuli used here were the cross spliced set created for the purpose of this experiment. Two conditions were included: a natural stimuli condition and a modified duration condition and 45 minimal pairs were presented in each condition. Participants were instructed to respond by pressing relevant keys on the computer keyboard indicating whether the third word on each trial was same as the first word or same as the second word in the auditory sequence.

Perceptual Identification (PI) Task and auditory Discrimination (AB-X) task with word stimuli (control task): PI and AB-X Discrimination tasks were also administered as an exact repetition of those tasks described in Chapter 2 (again, using natural and modified duration auditory stimuli arranged in two conditions), 45 minimal pairs were presented in each condition. However, cross spliced stimuli were not used in this case. The aim of these tasks was to serve as control tasks due to the fact that participants were different in the study of Chapter 2 and the study described here.

3.2.3 Results

Figure 3.6 and Table 3.3 show accuracy scores across all participants for the perceptual identification (PI) tasks (i.e. PI Cross Splice task and PI Control task) for the natural and modified duration condition. There were ceiling effects for English child and adult groups across tasks and task conditions. Greek adults did better than Greek children overall, with both groups being near chance levels for both PI task conditions using modified duration stimuli, unlike natural duration stimuli conditions. This suggests that duration cues that are not available with the modified duration condition affect performance, thus Greek participants (adult and child groups) use duration as a primary cue across all PI task types.

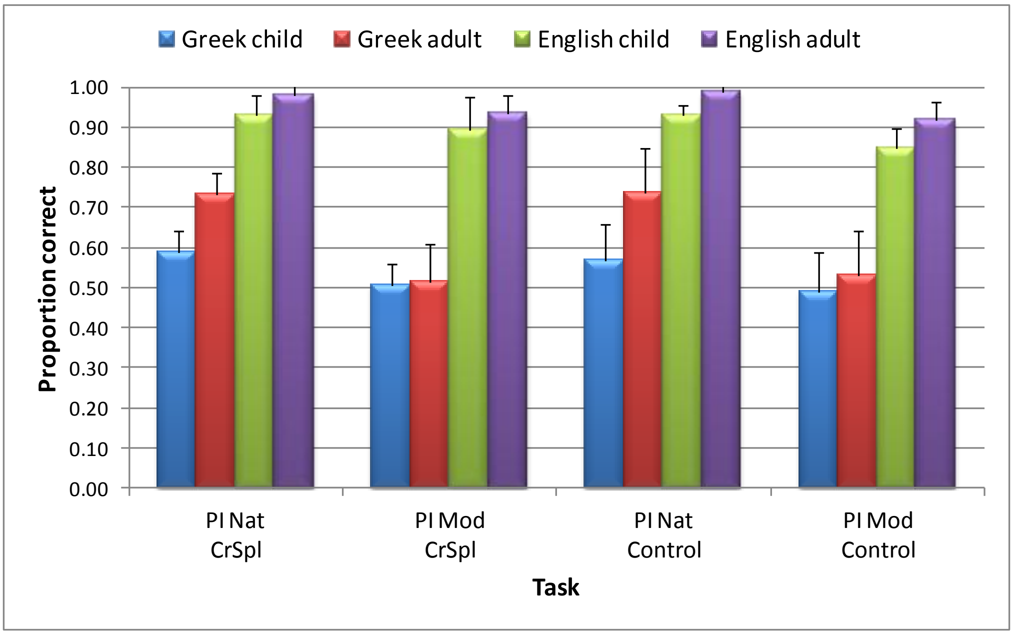


Figure 3.6: Accuracy scores for adult and child groups for Perceptual Identification (PI) Cross-Spliced (CrSpl) and control tasks for Natural (Nat) and Modified (Mod) duration condition.

Table 3.3: Mean accuracy scores and SD for adult and child groups for Perceptual Identification (PI) Cross-Spliced (CrSPL) and control tasks for Natural (Nat) and Modified (Mod) duration condition.

Group	PI Nat CrSpl n=45		PI Mod Crspl n=45		PI Nat Control n=45		PI Mod Control n=45	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek child (n=10)	0.58	0.09	0.51	0.09	0.56	0.12	0.48	0.13
Greek adult (n=10)	0.73	0.11	0.5	0.11	0.73	0.08	0.53	0.12
English child (n=10)	0.93	0.02	0.89	0.04	0.93	0.07	0.85	0.04
English adult (n=10)	0.98	0.01	0.93	0.04	0.98	0.01	0.92	0.05
Mean	0.81	0.05	0.71	0.07	0.81	0.07	0.69	0.08

There was a main effect of Language ($F_{(1. 36)}$ = 258.273, $p< .001$ and $F_{(1. 36)}$ = 179.304, $p< .001$, for the Cross Spliced and Control tasks respectively, see Figure 3.6 and Table 3.3), where English groups performed significantly better than Greek groups. There was a main effect of Age for both Cross Spliced ($F_{(1. 36)}$ = 8.283, $p< .05$) and Control tasks ($F_{(1. 36)}$ = 10.425, $p< .05$), since adults performed significantly better than children (see Figure 3.6 and Table 3.3). There was also a main effect of Vowel-length for both Cross Spliced ($F_{(1. 36)}$ = 74.665, $p< .001$) and Control tasks ($F_{(1. 36)}$ = 57.150, $p< .001$), since performance on the natural vowel duration was better than performance on the modified duration (see Figure 3.6 and Table 3.3). Also, the Vowel-length x Language interaction gave significant results ($F_{(1. 36)}$ = 23.189, $p< .001$ and $F_{(1. 36)}$ = 4.339, $p< .05$, for Cross Spliced and Control tasks respectively), because English participants showed a smaller difference between the two vowel length conditions (natural vs. modified duration) whereas Greek participants showed a larger difference between the two vowel length conditions (natural and modified duration) (see Figure 3.6 and Table 3.3). The Vowel-length x Age interaction was significant ($F_{(1. 36)}$ = 10.500, $p< .05$ and $F_{(1. 36)}$ = 4.770, $p< .05$, for Cross Spliced and Control tasks respectively), since the adult participants showed a larger difference between the two conditions (natural and modified duration) compared to child participants (see Figure 3.6 and Table 3.3). Similarly, the Vowel-length x Age x Language interaction was significant in both PI tasks, ($F_{(1. 36)}$ = 8.039, $p< .05$ and $F_{(1. 36)}$ = 4.624, $p< .05$, for Cross Spliced and Control tasks respectively). No other interaction effects were significant.

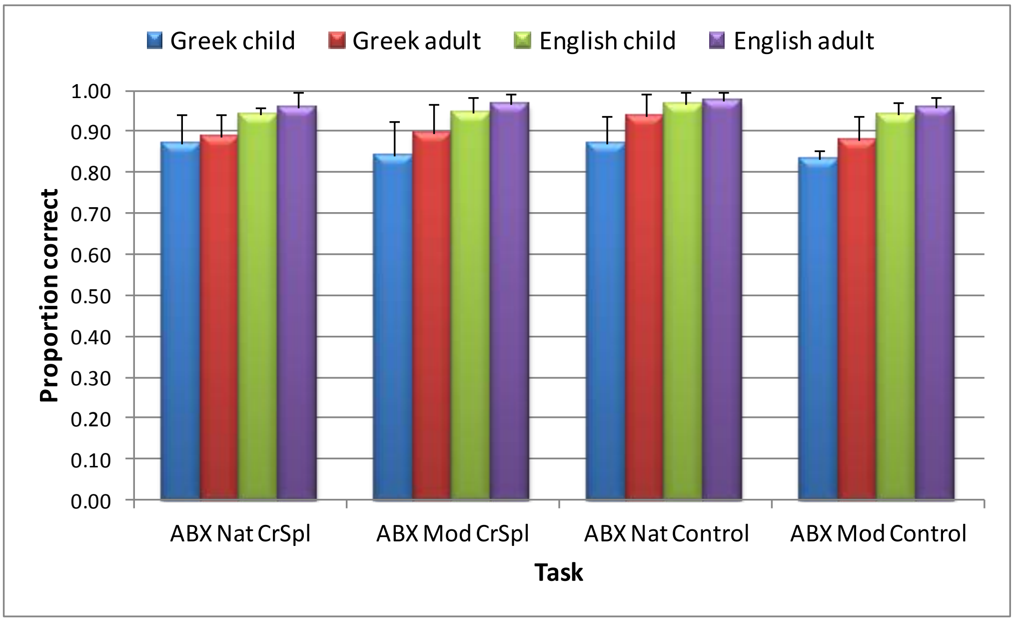


Figure 3.7: Accuracy scores for adult and child groups for all AB-X Discrimination tasks for Natural (Nat) and Modified (Mod) duration condition.

Table 3.4: Mean accuracy scores and SD for adult and child groups for Discrimination (ABX) Cross-Spliced (CrSPL) and control tasks for Natural (Nat) and Modified (Mod) duration condition.

Group	ABX Nat CrSpl n=45		ABX Mod Crspl n=45		ABX Nat Control n=45		ABX Mod Control n=45	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek child (n=10)	0.87	0.06	0.84	0.08	0.86	0.06	0.83	0.02
Greek adult (n=10)	0.88	0.05	0.89	0.06	0.93	0.05	0.87	0.06
English child (n=10)	0.94	0.02	0.94	0.04	0.96	0.02	0.94	0.02
English adult (n=10)	0.96	0.03	0.96	0.02	0.97	0.02	0.95	0.02
Mean	0.91	0.04	0.91	0.05	0.93	0.03	0.89	0.03

High accuracy scores were observed overall for the AB-X Discrimination tasks for all conditions and across participant groups, unlike the PI tasks. There was a main effect of Language ($F_{(1. 36)} = 4.894$, $p< .001$ and $F_{(1. 36)} = 16.128$, $p< .001$, for the Cross Splice and Control conditions respectively, see Figure 3.7 and Table 3.4) since the English groups performed significantly better than the Greek groups. There was also a main effect of Age ($F_{(1. 36)} = 5.650$, $p< .05$, for the ABX Cross Splice condition) since the adults performed better than children for this task. There was a main effect of Vowel-length for both Cross Spliced ($F_{(1. 36)}$ = 19.381, p< .001) and Control tasks ($F_{(1. 36)}$ = 26.401, p< .001). This confirms that duration affected performance for both tasks, where conditions with natural stimuli had higher scores compared with modified stimuli. All other two-way interactions were not significant. However, the Vowel-length x Age x Language interaction was significant for the ABX Cross Splice task ($F_{(1. 36)}$ = 9.119, p< .05), since as shown in Figure 3.7 and Table 3.4, the Greek adults performed considerably better than the Greek children in the natural duration condition than the modified duration condition, whereas English participants showed a similar pattern of performance between the two conditions.

3.2.4 Discussion

This experiment aimed to test the possibility that the consonants surrounding the vowel in each word stimulus may affect perceptual identification or discrimination as opposed to a perceptual choice being made upon the vowel alone. Two task types were administered, perceptual identification and discrimination tasks. Both included natural and modified duration stimuli type, based on vowel duration. However, cross-spliced stimuli were also used where the C_C environment of each minimal pair counterpart was cross-spliced with the other. This aimed to investigate whether the surrounding consonants offer 'cues' that perceptually aid identification and discrimination of the vowel which is the target speech sound across the experimental tasks. By cross-splicing the surrounding consonants, any potential 'cues' would confuse rather than aid participants for correct perceptual identification and discrimination scores.

The similar accuracy levels of each group for both cross-spliced and control tasks (PI and AB-X alike) suggest that there is no interference between the vowel under consideration and the surrounding consonants in the case of perceptual identification and discrimination of the vowel minimal pairs. Therefore, the preceding and following consonants do not carry necessary ‘cues’ that aid accuracy levels for the vowel perception. Results, however, in this experiment confirm the earlier assumptions of Experiment 1 and 2 that duration is used as a critical cue for vowel identification by the Greek learners of L2 English (adult and child groups) as performance was overall better for the natural condition compared to modified condition for the PI tasks.

The proposal by de Jong & Zawaydeh (2002) that consonants surrounding a vowel could generally affect the vowel quality is not confirmed in this case. It is important to note that the de Jong & Zawaydeh (2002) study focused investigations on whether identity of the final consonant (voiced or voiceless) affects vowel quality in Arabic which is a slightly different paradigm to what was used here. They report how durational correlates of phonemic contrasts are affected by stress and focus and specifically examine how quantity differences and voicing affect vowel durations and formants. They also contrasted vowels in stressed and unstressed syllables in the form of minimal pairs. An interesting finding they discuss is that generally segmental focus on voicing increase vowel durations. The experiment reported here aimed at examining de Jong & Zawaydeh (2002)’s observation hypothesizing that final consonants could affect vowel quality, and also extending this hypothesis to initial consonant position. Within the context of this experiement the target language was English (L2) and the paradigm used cross-spliced minimal pairs for both initial and final consonants. Observation of results reported here in view of the de Jong & Zawaydeh (2002) study could suggest that even if vowel quality may still be affected to some extent by the consonant environment in terms of other features (e.g. consonant voicing, place and manner of articulation, co-articulation features in a CVC environment, and more, e.g. Jacewicz & Fox, 2008; de Jong, 2004; de Jong & Zawaydeh, 2002, although the effect of these features was not tested in the current study), the necessary vowel dimensions toward accurate vowel perceptual identification and discrimination are still preserved in the case of cross-spliced stimuli.

3.3 General Discussion

The two experiments described in this chapter aimed at examining perceptual identification and discrimination accuracy of phonetic categories through the use of picture and cross spliced stimuli. Picture stimuli replaced previous word stimuli presented visually on the computer screen for participants. The aim was to explore the possibility of pictures conveying less 'cues' compared to word stimuli, or the possibility of word stimuli conveying 'cues' that aid perceptual identification, such as the number of letters making up each word (or representing each vowel sound) of a minimal pair. It was hypothesized that those minimal pair counterparts that contain a long vowel could be chosen in the case of a long-sounding vowel (in the auditory stimulus) based on the number of letters that word consists of (i.e. more letters), whereas a short-sounding vowel could lead to the selection of the 'shorter' word in terms of the 'less' number of letters consisting the word. Therefore, if the above hypothesis is true then in case the visual stimulus is replaced by picture stimuli instead of word stimuli, performance could drop.

Cross spliced stimuli were used as a different condition and both perceptual identification and auditory discrimination tasks were performed. The consonant 'frame' of the auditory minimal pair stimulus had been cross spliced; that is, the vowel of each word in the minimal pair had been digitally moved to replace its vowel counterpart. Therefore, if the consonants of each minimal pair word carry important information or 'cues' that could be affecting perceptual identification and discrimination of the vowel constituent then this would be reflected in the results. A 'control' set of tasks was also performed for comparison and in order to test any task type effects.

First, findings suggest that Greek participants (both adult and child groups) rely on duration as a primary cue rather than spectral cues that are present. This was observed across all task types based on the task conditions used (i.e. natural and modified duration). Performance was lower for Greek groups for all tasks and the modified duration condition in particular. This type of pattern is suggestive of heavy reliance upon duration as a primary cue. It should be noted that this effect was larger for the perceptual identification tasks compared to the auditory discrimination tasks revealing that Greek participants are able to discriminate the auditory differences of vowel minimal pairs when presented comparatively (in the AB-X format) but other task types cause confusions. This, however, was not the case with English control groups that show ceiling effects across all task types. It could therefore be suggested that Greek participants primarily use the 'wrong' cues for perceptually identifying English vowel distinctions and performance drops in the case these cues are not available, irrelevant or misleading. This is also in line with the assumptions of Chapter 2.

Second, the PI Picture task that used pictures as visual stimuli showed that Greek participants scored lower compared with the PI control task where visual stimuli were words even though auditory stimuli were identical in both tasks. The ceiling effects observed from English control groups for these tasks as well as the high accuracy scores by the Greek groups in the paper-based Picture Match and Translation tasks exhibit no doubt that the picture stimuli used could be problematic. However, it could suggest that there may be a 'link' between orthography and perceptual identification serving as an additional cue for L2 speakers. It is important to note here that results can only be suggestive rather than exhaustive. One reason is the limited amount of stimuli that were used in the picture stimuli task. This hypothesis could be further tested by using more minimal pair stimuli (auditory and visual) as well as a wider number of vowel contrasts.

Third, cross spliced stimuli do not seem to affect perceptually the identification or discrimination of the vowels under consideration. Performance was at similar levels as the control PI and AB-X discrimination tasks and followed the same pattern between the two conditions with higher scores for the natural duration condition and lower scores for the modified duration condition suggesting that the surrounding consonants do not appear to affect performance or offer any additional cues for the vowel perception than the vowel itself.

Fourth, task type seems to affect results with picture stimuli causing more confusion compared to word stimuli and PI tasks showing lower accuracy scores for Greek learners of L2 English than AB-X discrimination tasks compared to English control groups. This suggests that it is easier for Greek groups (child and adult participants alike) to discriminate minimal pair sounds. However, when comparison is not accessible between the two words (in the case of PI tasks) duration is used as a primary cue as data from natural and modified duration conditions reveal.

Finally, L2 child participants' low scores in the perceptual identification task compared with L2 adult participants' scores could be attributed to a combination of age and proficiency effects. The fact that L2 child participants had higher accuracy scores for the auditory discrimination tasks and the paper-based (Picture Match and Translation) tasks compared with their lower scores (at chance level) for the perceptual identification task types indicates task bias rather than lexical bias. Age and proficiency level could, however, be contributing factors which should be further tested in future experimental designs in order to identify which of these factors is more likely. Therefore, age effects that are observed for L2 child versus adult participants could simply be interfering with differences in proficiency levels. The current experimental design would not allow for clarifying the contribution of either factor (age versus proficiency) for the two age groups which could be suggested to have implications in terms of how current results are interpreted, particularly in relation to L2 child participants. Experiment 4 (Chapter 4) discusses this issue from a perceptual training angle which aims to shed some light to proficiency level effects in relation to age.

Chapter 4

Perceptual Training of phonetic categories: High-variability auditory training for Greek child and adult L2 English learners.

This chapter[28] presents empirical work from a perceptual training experiment, completed by child and adult Greek speakers of L2 English, which aims to demonstrate whether specialised training could allow the L2 learners for correct cue weighting in perceptual identification and discrimination of L2 phonetic segments by shifting perception into relevant and critical cues, or reallocating attention to those cues that are critical. Perceptual improvement was examined as a result of training using the 'high-variability phonetic training' (HVPT) technique (Logan *et al*., 1991; Yamada, 1993; Pisoni *et al*., 1994) used successfully in several previous studies to improve performance of phonemic identification (e.g. Bradlow *et al*., 1997, 1999; Lively *et al*., 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Pisoni *et al*., 1994; Iverson, Hazan, & Bannister, 2005 ; Uther *et al*., 2007; Iverson & Evans, 2009; Ylinen *et al*., 2010, see Chapter 1 for a review of HVPT).

The term 'high variability' refers to the use of stimuli from multiple speakers and multiple phonetic contexts. High-variability phonetic training technique involves multiple exemplars of different words pronounced by multiple speakers, both male and female. It features speech sound contrasts (e.g. 'lake' vs. 'rake') of a range of linguistic variability aimed at resembling the context variability normally encountered by L2 learners and is known to help improve

[28] Part of the work comprising this chapter has appeared in Giannakopoulou *et al*. (2011).

the identification and discrimination of non native speech contrasts (Logan *et al.*, 1991; Lively *et al.*, 1993, 1994; Iverson *et al.*, 2005; see also Ylinen *et al.*, 2010). This variability is essentially used to aid the L2 learner to identify the perceptual constancy for a range of linguistic contexts (Uther *et al.*, 2005). Previous research has emphasised the significant role of acoustic-phonetic talker characteristics in speech perception and speech processing, particularly for L2 learners (e.g. Bradlow, Tarretta, & Pisoni, 1996).

The range of different vowel inventories in different languages (ranging between three-vowel to fifteen-vowel systems) poses different degrees of difficulty for L2 vowel category perception (and production) for L2 learners (Iverson & Evans, 2007a), as already discussed in Chapter 2. Such learning constraints are often attributed to how L2 learners weight cues for perceptual identification and discrimination of L2 speech sounds (e.g. Escudero, 2000; Cebrian, 2006). HVPT allows L2 learners to be exposed to a broad range of speech stimuli, as well as multiple speakers, in a number of consecutive sessions. Therefore this high-variability training aims at the reallocation of attention to those cues that are critical for correct perceptual identification and discrimination of L2 speech sounds. Similarly, by re-tuning the participants' attention to the critical cues, attention is shifted from those cues that are not relevant for speech sound recognition. The HVPT approach could be interpreted in view of the 'attention to dimension' (or 'A2D') models of speech perception (e.g. Iverson & Kuhl, 1995, 1996; Goldstone, 1993, 1994; McClelland, 2001; Nosofsky, 1986), based on a shared assumption that perceptual learning is dependent on specific changes in the distribution of attention, or reallocating the L2 learner's attention on the 'relevant' dimension. It is proposed that in order for new phonetic categories to be learnt, perceptual dimensions that are focal should be 'stretched' while those non-focal dimensions should be 'shrunk' (e.g. Francis & Nusbaum, 2002). Perceptual training is therefore aiming to shape the L2 learner's perception so that attention is shifted toward dimensions (or cues) that are relevant (e.g. Francis & Nusbaum, 2002; Kuhl & Iverson, 1995; Jusczyk, 1989).

Previous studies that have used high-variability training have shown that it not only helps L2 learners improve their perceptual identification of L2 speech sounds, but this improvement is also retained for a minimum of 3 months (shown by repetition tests within a 3-month period; Bradlow *et al*., 1999; Lively *et al*., 1994; Iverson & Evans, 2009), and it can be transferred to the L2 learners' production of L2 speech (Bradlow *et al*., 1997, 1999; Hazan *et al*., 2005; Lambacher *et al*., 2005; Lengeris, 2008). A variety of training techniques that have been conducted to date have shown that it is possible for L2 learners' performance to significantly improve perception (and production) of L2 sounds (e.g. Jamieson & Morosan, 1989; Logan & Pruitt, 1995; Iverson & Evans, 2007a; Nishi & Kewley-Port, 2008; Ylinen *et al*., 2009).

A number of studies have reported perceptual improvement for consonants (e.g., Logan *et al*., 1991; Lively *et al*., 1993; Lambacher, Martens, & Kakehi, 2002; Iverson *et al*., 2005; Pruit *et al*., 2006; Hazan *et al*., 2005) and vowels (e.g., Nishi & Kewley-Port, 2007, 2008; Iverson and Evans, 2009). Also suprasegmental features were used for training which resulted in the improvement of the perception of tone contrasts (e.g. Wang *et al*., 2003). An abundance of studies have investigated the English /r/-/l/ contrast in the case of Japanese adult learners of L2 English (e.g. Logan *et al*., 1991; Lively *et al*., 1993, 1994; Hazan *et al*., 2005; Lim & Holt, 2011) with the aim to reallocate the learners' attention to critical cues of the L2 contrast by means of the Perceptual Fading technique using enhanced contrastive stimuli. Other training techniques have involved cue manipulation such as the All Enhancement and Secondary Cue Variability method (e.g. Iverson *et al*., 2005). Audiovisual (AV) types of HVPT are also effective training techniques (e.g. Hardison, 2003; Ortega-Llebaria, Faulker, & Hazan, 2001, Hazan *et al*., 2005), based on the integration of visual and auditory cues for native speakers as shown in the McGurk effect (McGurk & MacDonald, 1976), where visual and auditory information are conflicting. For example, when visual presentation is /gi/ and auditory presentation is /bi/, the resulting sound perceived is usually /di/ (Munhall, Gribble, Sacco, & Ward, 1996), which suggests the importance of visual information for speech intelligibility. This is especially true in the case of difficult listening conditions due to environmental noise (e.g. Sumby & Pollack, 1954). A recent extension to this is a combination of auditory and articulatory training suggesting that both training interventions can in fact improve correct cue weighting for Spanish/Catalan learners of L2 English (i.e. Aliaga-Garcia, 2011). The HVPT technique overall appears to be a robust and reliable method for L2 learning.

Lengeris & Hazan (2010) in a recent study used perceptual training with Greek speakers of L2 English, examining perception and production of L2 sounds as well as investigating individual differences in learning and task performance. The study examined how individual L1 vowel processing ability and frequency discrimination acuity could affect the learning of L2 vowel categories. Speech and non-speech categories were included in the experimental conditions. Although the role of L1 interference was not rejected (at least at group level) as it is supported by a variety of L2 learning models (e.g. Best, 1995a; Best *et al.*, 2001; Kuhl, 1991; Kuhl *et al.*, 1992; Kuhl & Iverson, 1995; Flege, 1995a, 2003; Iverson *et al.*, 2003), the data suggest that the degree of L1 interference or general L1 biases was highly dependent on the individual as some are better at processing spectral and acoustic information at both the start and end points of a training intervention. In particular, individuals with better frequency discrimination for synthetic vowels in L1 and L2 and nonspeech stimuli were generally also better at identifying natural L2 vowels before and after training. Because most studies concentrate on group performance and overall training trends, examining individual differences is an important part of determining the success of perceptual training as well as general L1 – L2 speech processing (see also, Iverson, Ekanayake, Hamann, Sennema, & Evans, 2008; Hattori & Iverson, 2009; Iverson & Evans, 2007b, 2009).

L2 phonetic cue weighting was tested using the English tense /iː/ versus lax /ɪ/ distinction in the case of Greek learners of L2 English. Two different conditions were used in the current training study: training with natural duration stimuli and training with modified duration stimuli (ambiguous or equalized duration cues; see section 4.1.2 for more detail on the stimuli used). Stimuli consisted of a list of minimal pairs. One set of stimuli included natural duration auditory stimuli where duration and spectral cues were available for perceptual identification and discrimination tasks. A second stimuli set included modified duration auditory stimuli which were manipulated in a way that the use of the spectral cues was ‘forced’ for perceptual identification and discrimination. This means that the categorization of phonemic categories was not possible solely on the basis of the non-critical (duration) cues, but purely on spectral cues. Given that duration is not phonemically relevant within the Greek vowel inventory, the aim was to explore whether there was any effect of removing these cues and how high variability training could improve cue weighting and performance.

This study explores whether learnt L1 cue weighting affects how cues are weighted in the L2 in the case of Greek learners of L2 English and whether specialised perceptual training, used as an intervention, could allow the L2 learner for correct cue weighting in perceptual identification and discrimination of L2 phonetic segments, thereby re-arranging perception into relevant and critical cues. The first aim of this study was to test performance and compare identification and discrimination task results between Greek learners of L2 English and English native speakers, for child and adult participants alike. The second aim was to determine whether high-variability phonetic training (HVPT) could have an effect on L2 cue weighting for the Greek learners of L2 English. The third aim was to explore whether stimulus type (i.e. natural vs. modified duration) may have any effect on perceptual training in how Greek participants learn to weight primary versus non primary cues, thus resembling a more native-like weighting. Given that there are several theories of speech acquisition that emphasize the role of age of acquisition (e.g. Kuhl *et al*., 2008; Flege, 2002), it was thought useful also to explore whether there were maturational differences in acquisition of new L2 phonetic categories. The performance of L2 learners of different age groups was therefore examined in order to determine whether there were similar patterns of difficulty in adult versus child learners. Although both adult and child participants have previously taken part in perceptual identification experiments, this is the first HVPT study, to the author’s knowledge, that involved child participants.

This study uses a 'generalisation' training protocol where participants are trained with fewer word pairs and tested with more word pairs. Specifically, training includes more talkers and fewer words, while the pre and post training test includes trained and untrained words pronounced by an unfamiliar talker which would allow to test the generalization of learning. Nishi & Kewley-Port (2007a, see a discussion in Section 1.4.2) have demonstrated significant perceptual improvement for the /r – l/ contrast by Japanese learners of L2 English using a similar training protocol. This study aims to show whether such outcomes would also be true for Greek adult learners of L2 English, and also extend it to Greek child learners of L2 English.

Also, child participants chosen for this study were early learners of L2 English, aged 7-8 years. This age criterion stemmed out of experimental results of Experiment 1 (Chapter 2) demonstrating that the younger child group (aged 7-8 years) performed worse across all experimental tasks compared with older child groups (9-10 and 11-13 years) and adults. This 'age effect' was previously discussed in view of the low proficiency level of the 7-8 year old group due to the limited amount of L2 instruction received in comparison with the other L2 groups at the time of testing. This experiment, however, aims to investigate whether factors such as proficiency level and L2 experience that have been considered to interact with age in previous experiments, could be reversed due to the learning capacity of younger participants. As proposed by Flege (1992), while the L1 phonetic system develops and stabilises in early and middle childhood, it is more difficult to establish new phonetic categories. It would therefore be necessary to test L2 learners as early as possible in order to test whether new phonetic categories could be established (or at least start being established) following targeted phonetic training (see also, Perani *et al*., 2003; Neville *et al.*, 1997; Neville *et al.*, 1992) and what the outcomes may be in relation to adult L2 learners who already have greater L2 experience yet a more limited learning capacity (e.g. Long, 1990; Chiswick & Miller, 2008).

Two main research questions are discussed in this study: 1. Whether there is plasticity in perceptual learning for child and adult L2 learners alike, as well as possible age differences between child and adult learners; 2. What are the benefits of the HVPT technique for L2 speech perception. Previous studies (e.g. Bradlow *et al*., 1997, 1999; Lively *et al*., 1993; Lively *et al.*, 1994; Logan *et al*., 1991; Yamada, 1993; Pisoni *et al*., 1994; Iverson *et al.*, 2005; Nishi & Kewley-Port 2007a; Iverson & Evans, 2009; Ylinen *et al*., 2010) allow the prediction that there will be plasticity in perceptual learning for L2 learners, even though the degree of perceptual improvement is difficult to accurately predict, especially for child L2 learners. Also, previous research predicts that the HVPT technique for L2 speech perception is beneficial for adult L2 learners. It is hypothesized that it will also be beneficial for child L2 learners. Given the learning capacity of children, as discussed in Chapter 1, it would be an interesting investigation to see whether it is adult or child L2 learners who will benefit more from the HVPT technique.

4.1 Method

4.1.1 Participants

Adult groups: Twenty native adult speakers of Standard Modern Greek (13 female, 7 male) aged 20-30 (mean age = 27.4) were tested. They had all lived in Greece and had all studied English as L2 in school (public and private education; L2 English education ranged 8-9 years; mean = 8.7 years). Their English proficiency was self-rated as ‘advanced’ (on a scale of basic, intermediate and advanced) in listening, speaking, reading and writing (and all had received the Cambridge Certificate in Advanced English or Certificate of Proficiency in English) as recorded in the language background questionnaire that all participants completed before testing. They were recruited from a private language school in Greece. None had spent more than 2 weeks in an English-speaking environment. Twenty monolingual English native speakers (14 female, 6 male) aged 19-30 (mean age = 21.4) were also tested and served as controls and were recruited from Brunel University. English native speakers received course credit for their participation. Greek learners of L2 English received a small gift of stationery items as reward for their participation and were entered into a prize draw. All adult participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

Child groups: Twenty child native speakers of Standard Modern Greek (10 female, 10 male), aged 7-8 (mean age = 7.9) were tested. They were recruited from a private language school in Greece. They had all lived in Greece and were students of L2 English (public and private education). Their self-rated level of proficiency in English was 'basic' (on a scale of basic, intermediate and advanced) in listening, speaking, reading and writing as recorded in the language background questionnaire that all participants completed before testing (L2 English education ranged 1-2 years; mean = 1.4 years). None had spent more than 2 weeks in an English-speaking environment. Twenty child monolingual native speakers of Standard English were also tested and served as controls (14 female, 6 male; mean age = 7.6). They were recruited from a local primary school in the UK. All child participants received a small gift of stationery items as reward for their participation. Greek child participants were entered into a prize draw. All child participants had normal or corrected to normal vision and none reported any history of a speech or hearing impairment.

4.1.2 Stimuli

4.1.2.1 Pre and Post test stimuli

Forty-five minimal pairs (i.e. word pairs that are differentiated by one single vowel sound, e.g. *sit* versus *seat*) of the English /iː/ versus /ɪ/ (tense - lax) vowel distinction were used in total for the pre and post training tests (thus, 45 visual stimuli arranged as minimal pairs, e.g. *beat – bit*, and the respective 90 digitally recorded words were used as auditory stimuli). The auditory stimuli[29,30] were words pronounced by a male native speaker of English

[29] The same stimuli have been previously used in Ylinen *et al.* (2010).
[30] The stimuli used in this experiment are the same as described in Section 2.2.1 (see also Appendix 3). For a detailed account on the recording procedure, refer to Chapter 2.

representing typical Southern British English pronunciation (37 years of age). All auditory stimuli were natural speech items with normal vowel durations.

A second set of 45 minimal pairs was also used, where the normal vowel duration was digitally manipulated with Praat software (Boersma & Weenink, 2004) so that for each minimal pair the intrinsically long /iː/ vowel was shortened to match the duration of the intrinsically shorter /ɪ/ vowel and vice versa (see Figure 4.1). This was the same modified duration set also used in the experimental studies described in Chapter 2 and Chapter 3. For example, for the minimal pair *fit* – *feet*, the sound stimulus would be either [fiːt] as in *feet* but the long vowel /iː/ was shortened to match the duration of the short vowel /ɪ/ or [fɪːt] as in *fit* but the short vowel [ɪ] was lengthened to match the duration of the long vowel /iː/. Vowel modifications[31] were resynthesized with the Pitch-Synchronous Overlap and Add (PSOLA) technique (Boersma & Weenink 2004). In these modified stimuli, the spectral quality of vowels was preserved even though duration cues were ambiguous (thus not available) (see Figure 4.1) which forced participants attend to spectral cues only.

Thus, 90 digitally recorded words[32] arranged as 45 minimal pairs were used as auditory stimuli during the pre and post training test and the respective 45 minimal word pairs (orthographic form) were used as visual stimuli. The auditory manipulation that resulted in the modified duration words produced a second set of 45 auditory minimal pairs while visual stimuli were not altered or modified in any way.

[31] A detailed account on the vowel modification procedure is provided in Chapter 2.
[32] All auditory and visual stimuli used in this experiment have also been previously used in experimental tasks described in Ylinen *et al.* (2010).

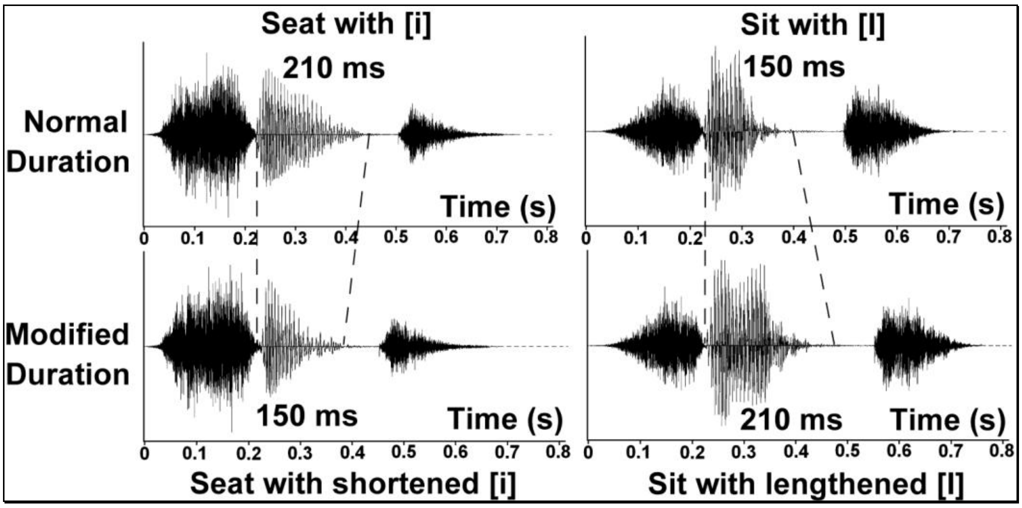


Figure 4.1: An example of the stimuli used. Top: The minimal pair ‘seat’ and ‘sit’ with normal duration (i.e., as ordinarily pronounced by a native English speaker). Bottom: The same minimal pair with modified vowel durations: [iː] was shortened to correspond to the original duration of [ɪ], and [ɪ] was lengthened to correspond to the original duration of [iː].

4.1.2.2 Training Stimuli

The auditory stimuli used in the training sessions were pronounced by 4 different native British English speakers (2 male and 2 female speakers) considered to represent typical Southern British English pronunciation. None of the speakers used in the training phase were used in the test phase. Nineteen minimal pairs out of the forty-five used in the pre- and post-training test, were used for the training sessions. Auditory stimuli were modified following the same procedure as described for the auditory stimuli used in the pre and post training tests (for natural and modified duration stimuli alike). Therefore, each stimulus had two versions: 1) a normal vowel duration, 2) a modified vowel duration. The 4 speakers used for the training stimuli were different to the speaker who pronounced the stimuli for the pre- and post-training tests.

4.1.3 Apparatus

Two laptop computers (AMD Sempron and Intel Core 2 Duo T5550) were used with E-Prime 1.1 (Schneider *et al*., 2002a,b) and a version of the HVPT programme (Logan *et al.*, 1991, Pisoni *et al.*, 1994) which allowed two participants to train simultaneously. High quality headphones (SONY MDR-V150) connected to each laptop computer, were used to present the auditory stimuli. Visual stimuli were displayed on a 33 x 20 cm monitor. Reaction times (RT) and responses were automatically recorded for each participant through the E-Run[33] software application of the E-Prime 1.1 (Schneider *et al*., 2002a,b) software for the pre and post training sessions. Participants responded by pressing an allowable relevant button on the computer keyboard (for the pre and post training test) or click a mouse button (for the training sessions) which triggered the next trial after a 1000ms delay. All auditory stimuli were binaurally presented at 44 kHz, 16-bit resolution and at a comfortable listening level (varying between 65-75 dB).

4.1.4 Procedure

This experiment involved three stages for the Greek (adult and child) participants: pre-training test, training, and post-training test. In the pre-training test session, the use of spectral and duration cues of the English /iː/ and /ɪ/ vowel distinction (e.g. *seat* vs. *sit*) was examined, comparing Greek and English speakers. The post-training test session (i.e. a repetition of the pre-training test) aimed at examining whether the training intervention had a significant effect in vowel perception and cue weighting, through comparison of the pre- and post-training test results. The pre- and post-training test sessions were identical. English native speakers (control group) participated only in the pre-training test for comparative reasons.

[33] The E-Run software application is the presentation component of E-Prime 1.1 enabling for millisecond precision of stimulus presentation, as well as synchronizations of stimuli (e.g. visual and/or auditory) and data collection.

4.1.4.1 Pre-training and Post-training tests

Two types of tasks were used to examine perceptual identification and discrimination (and consisted of the same identification and discrimination task-types as described in Chapter 2):

Perceptual Identification Task (PI): This task used 45 normal vowel duration and 45 modified vowel duration minimal pairs of English words arranged into two conditions (natural and modified duration stimuli). For each trial an auditory stimulus was presented through the headphones and a simultaneous visual stimulus (i.e. the orthographic form of the minimal word pair, e.g. *sit* – *seat*) was presented on the screen. For each condition there were 90 stimuli presented, 180 in total (45 minimal pairs x 2 tokens =90 stimuli x 2 conditions = 180 stimuli). For each participant, all stimuli were presented in random order arranged automatically by E-Prime (Schneider *et al.*, 2002a,b). The two counterparts (visual stimuli) of each minimal pair were also presented in random order. Participants were instructed to press a relevant key on the computer keyboard choosing the word they heard between the two words (minimal pair) presented on the screen. Due to the nature of the stimuli in the modified duration stimuli condition, participants could only base their choice on spectral rather than durational cues to make the vowel identification.

Auditory Discrimination Task (AB-X)**:** The stimuli were presented only aurally as a discrimination task, arranged in an ABX format (i.e. ABA or ABB). Each trial comprised a sequence of a minimal word pair (word A followed by word B or word B followed by word A) and a third stimulus being the exact repetition of either word A or word B. This task also included two conditions: a natural stimuli condition and a modified duration condition (with stimuli matched in duration to the other word in the minimal pair) and each condition consisted of 45 trials. Participants were instructed to respond by pressing relevant keys on the computer keyboard whether the third word on each trial was same as the first word or same as the second word in the auditory sequence.

4.1.4.2 Training sessions

The training program was based on the high-variability perceptual training procedures developed by Logan *et al.* (1991), extended by Yamada (1993) and others (e.g. Bradlow *et al*. 1997, 1999; Lively *et al*. 1993, 1994; Pisoni *et al*. 1994; Ylinen *et al*. 2010). The training phase consisted of 10 x 30 minute sessions (one session per weekday) in a 2-week period. Only Greek speakers participated in the training sessions. Like the pre-training testing paradigm, each trial consisted of an auditory stimulus (one of the words of a minimal pair) and the requirement to select a choice from a pair of visually presented minimal pair of words. Each training session had 304 (19 minimal pairs x 4 speakers x 4 presentations) trials (different combinations of speaker and word type) presented in random order. Each trial consisted of a single minimal word pair (each word of a minimal pair appearing on the left and right hand side of the screen respectively) and one of the two words was presented aurally. Each minimal word pair was repeated four times: twice in an AB format where the auditory stimulus corresponded with A and B respectively and twice in a BA format where the auditory stimulus corresponded with B and A respectively. Answers were triggered by clicking a computer mouse on the word selection displayed on the screen. Participants were instructed to select the word they heard at each trial by clicking the mouse on one of the two minimal pair words. The visual stimuli would remain on the screen until an answer was triggered. Unlike in the test sessions, the participants received feedback on each trial and had the option to replay the auditory stimulus if needed. 'Correction' trials were also included: these additional trials were given in the case of an incorrect response (with the visual positions of the minimal word pair randomly reassigned on the next trial to avoid guessing). The training program contained cartoon animations (happy vs. sad animation providing feedback on correct or incorrect responses) to motivate the children in particular. Every correct answer would also trigger a virtual yellow coin on the screen which aimed at being an indication for performance (i.e. the number of coins indicating the number of trials answered correctly). Figure 4.2 and 4.3 are screen shots of two random training trials representing positive and negative feedback respectively.

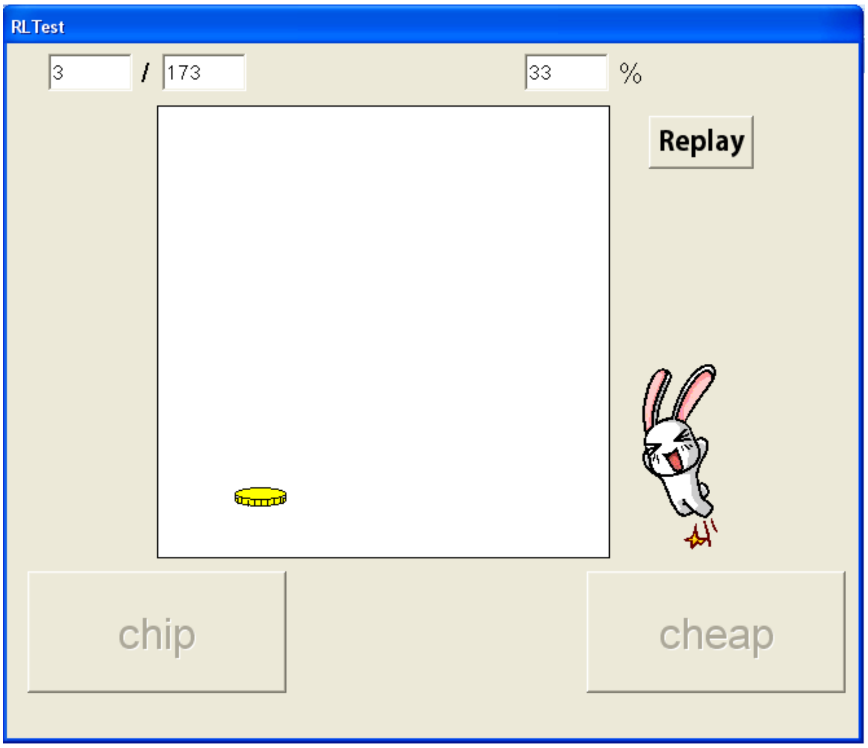


Figure 4.2: A screenshot from a training session where the selected answer was correct.

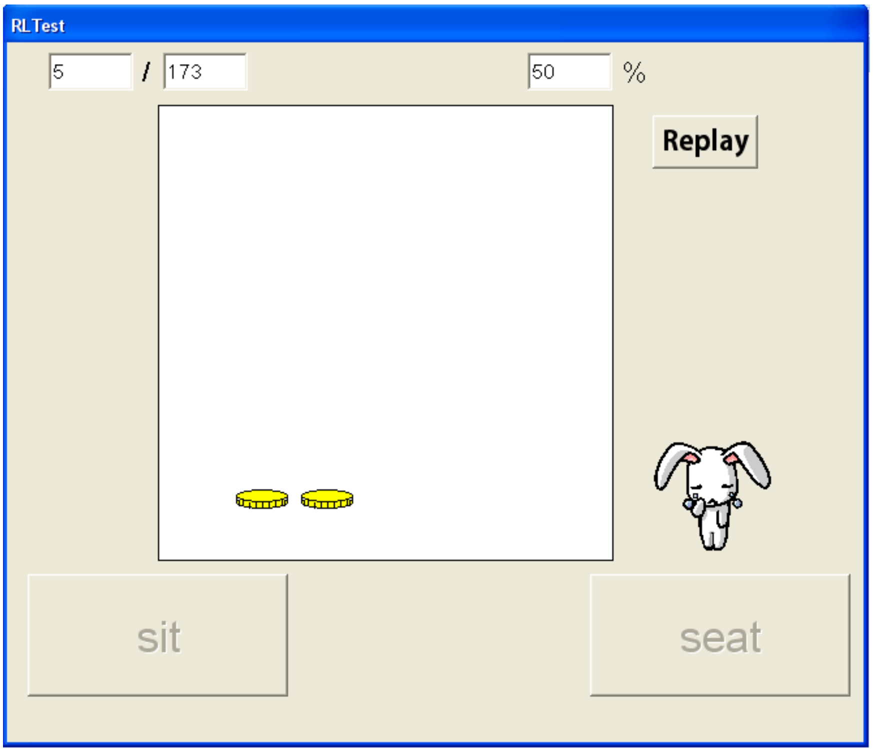


Figure 4.3: A screenshot from a training session where the selected answer was incorrect.

Training involved two conditions: natural vowel training and modified vowel training. Ten participants were trained using natural duration stimuli and ten participants were trained using modified duration stimuli. The reason was to investigate whether stimuli type (natural vs. modified duration) could affect perceptual training and to what degree. Participants completed a language background questionnaire, gave written consent and participated in all sessions. Participants were tested in convenient geographical locations. They were tested individually in sound attenuated cubicles.

4.2 Results

This section initially presents results from the pre-training test, comparing Greek and English child and adult groups in order to show the respective starting points (pre-training perceptual identification and discrimination level) for the Greek groups and these are compared with performance of the respective English groups. Post-training scores are then compared with pre-training scores for the Greek groups in order to demonstrate the effect of training. Difference scores are presented exploring which group (child or adult) benefited the most from the training intervention. Also, post-training results for the stimulus type used in the two different conditions aimed to examine whether cue-specific training could affect perceptual identification and/or discrimination of English vowel minimal pairs.

4.2.1 Pre-training results

For the pre-training results, correct responses were averaged across participants for the two language groups, age groups and task conditions. Accuracy scores were submitted to a three-way mixed design analysis of variance (ANOVA) with Language Group (Greek vs. English native speaker) and Age (child vs. adult) as between subject factors and vowel Length (natural vs. modified duration stimuli) as a within subject factor.

On the perceptual identification task, Greek participants performed worse compared with English speakers as revealed by a main effect of Language ($F_{(1,76)}$= 488.681, p< .001, see Figure 4.4 and Table 4.1). For the perceptual identification task, there was an effect of Age with adult participants performing better compared with child participants ($F_{(1,76)}$= 34.595, p< .001, see Figure 4.4 and Table 4.1). Identification of words with natural vowel length was better than those with modified duration vowels as revealed by a main effect of Vowel Length ($F_{(1,76)}$= 126.175, p<.001, see Figure 4.4 and Table 4.1), confirming the assumption that vowel duration was indeed used as a critical cue for perceptual identification performance. For child participants, there was less of a difference between natural and modified duration words compared with a larger difference between natural and modified duration words by adult participants. This was confirmed by a significant Vowel Length x Age interaction ($F_{(1,76)}$=17.856, p< .001, see Figure 4.4 and Table 4.1). There was a significant 3-way interaction ($F_{(1,76)}$=35.252, p< .001), indicating that in Greek speakers, the children showed less of a difference in performance in natural and modified duration conditions compared to the Greek adults. On the other hand, in the English native speakers, there was a consistent improvement in performance for the natural duration conditions. However, it is worth noting that in the Greek children, their performance was around chance level, so this difference is likely to be due to a floor effect (see Figures 4.5 and 4.6 below).

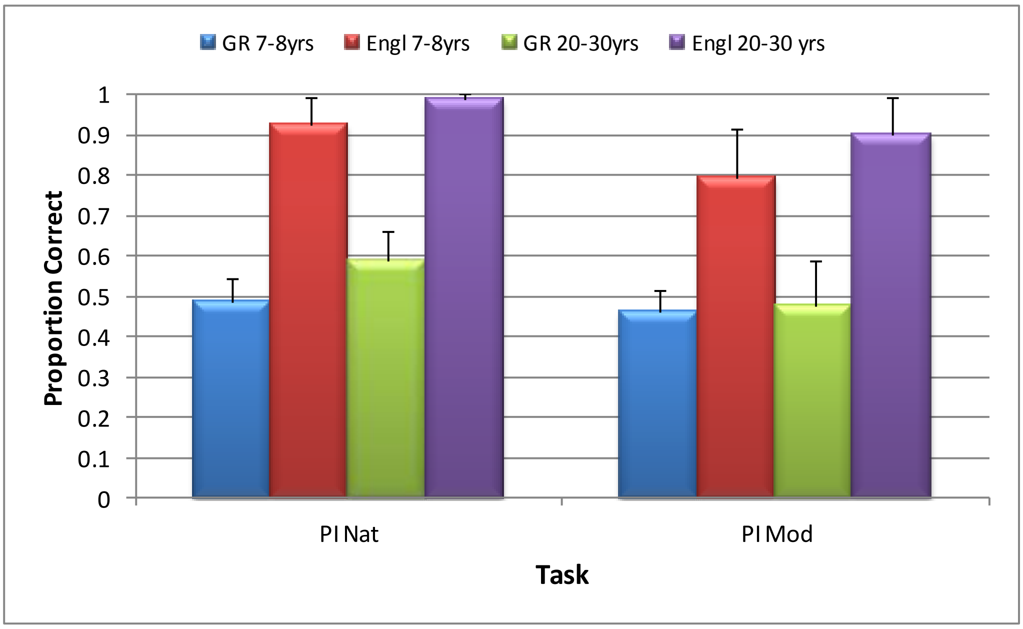


Figure 4.4: Accuracy scores for Greek child, Greek adult, English child and English adult participants for the Perceptual Identification (PI) task and for Natural (Nat) and Modified (Mod) duration conditions.

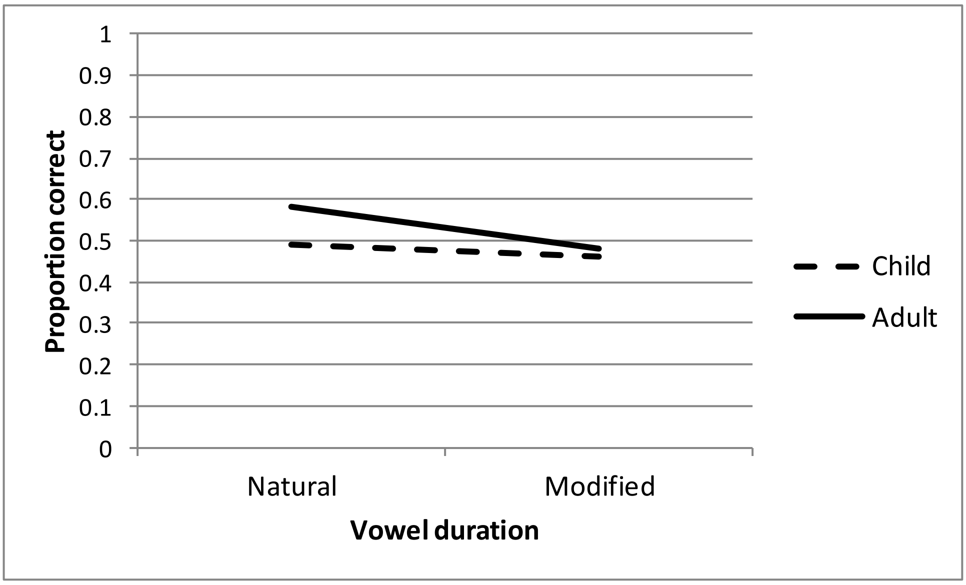


Figure 4.5: Accuracy scores for Greek speakers of both age groups on the Perceptual Identification Task.

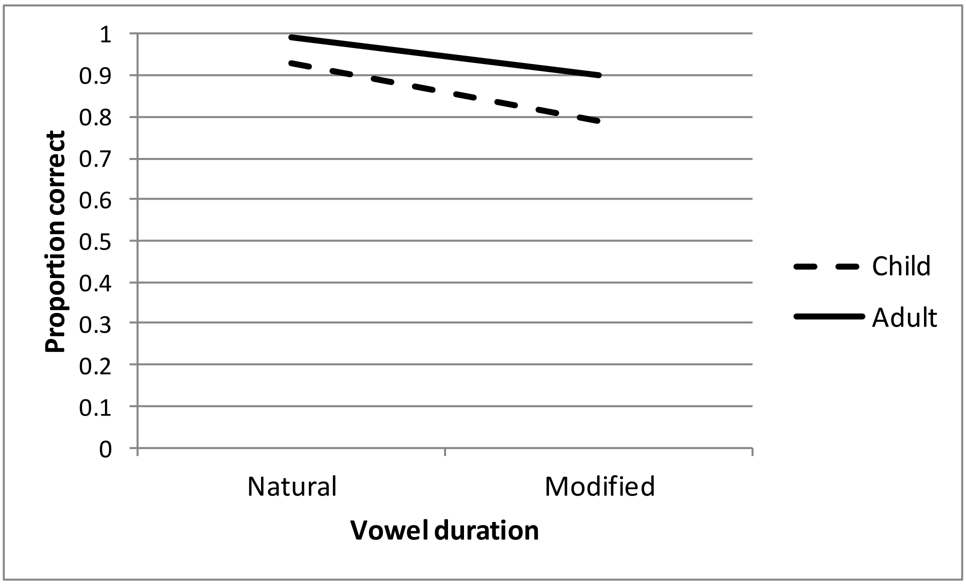


Figure 4.6: Accuracy scores for English speakers of both age groups on the Perceptual Identification Task.

Table 4.1: Mean accuracy scores and SD for adult and child groups for Perceptual Identification (PI) and Auditory Discrimination (AB-X) tasks for Natural (Nat) and Modified (Mod) duration condition.

Group	PI Nat n=45		PI Mod n=45		ABX Nat n=45		ABX Mod n=45	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek child (n=20)	0.49	0.06	0.46	0.05	0.76	0.10	0.72	0.11
English child (n=20)	0.93	0.07	0.79	0.12	0.97	0.03	0.95	0.03
Greek adult (n=20)	0.58	0.07	0.48	0.11	0.92	0.03	0.73	0.05
English adult (n=20)	0.99	0.02	0.90	0.10	0.93	0.07	0.95	0.05
Mean	0.78	0.05	0.66	0.10	0.89	0.06	0.84	0.06

For the AB-X discrimination task, English speakers performed better compared with Greek speakers as revealed by a significant main effect of Language ($F_{(1,76)}$=128.329, p< .001, see Figure 4.7 and Table 4.1). A main effect of Age also showed that adults performed better compared with children ($F_{(1,76)}$= 5.099, p< .05, see Figure 4.7 and Table 4.1). Discrimination of words with natural vowel duration was better compared with modified vowel duration words as revealed by a main effect of Vowel Length ($F_{(1,76)}$= 9.858, p<.05, see

Figure 4.7 and Table 4.1). There was a more marked difference between child and adult performance (with children considerably worse) in the Greek native speakers compared to the difference between child and adult in the English native speaking group, a pattern reflected in a significant Age x Language interaction ($F_{(1,76)}$=9.566, p< .05, see Figure 4.7 and Table 4.1). Greek groups also performed significantly better on the natural vowel duration condition relative to the modified duration condition whereas English groups did not show as large differences in performance between the two Vowel Length conditions and this effect was statistically significant as reflected by the vowel length by language interaction ($F_{(1,76)}$=12.613, p<.001, see Figure 4.7 and Table 4.1). The Vowel Length x Age x Language Group interaction ($F_{(1,76)}$=11.193, p< .001, see Figures 4.8 and 4.9 below) was significant, showing that as for the perceptual identification task, there was less of a difference in performance between modified and natural vowel duration conditions for the Greek child as compared to the Greek adult participants. By contrast, for the native English speakers, both groups were already near ceiling, and so there was very little difference between natural and modified duration conditions. No other main effects or interactions were significant.

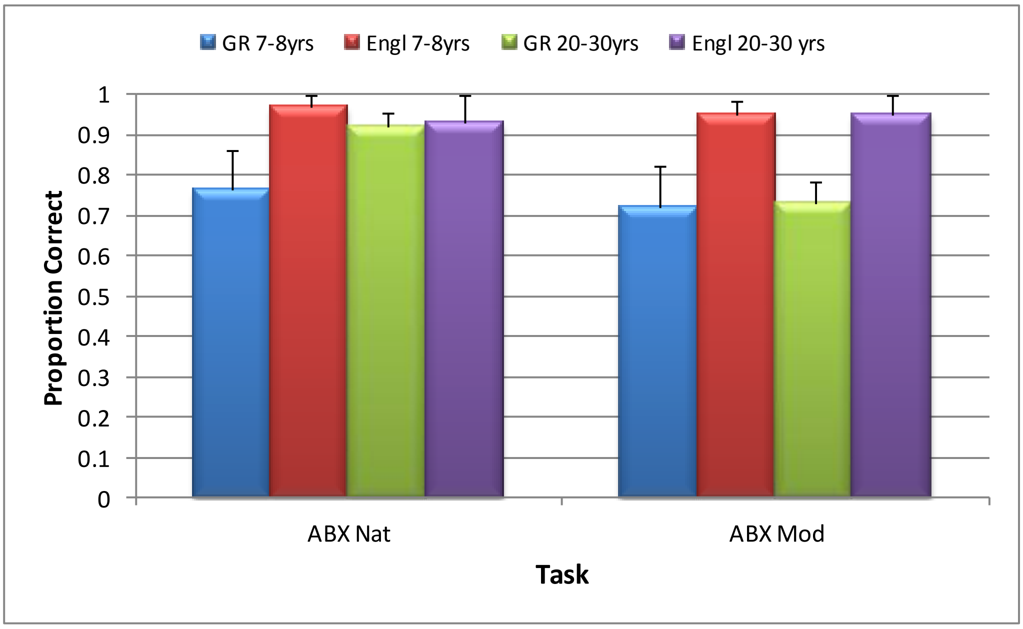


Figure 4.7: Accuracy scores for Greek child, Greek adult, English child and English adult participants for the Auditory Discrimination (ABX) task and for Natural (Nat) and Modified (Mod) duration conditions.

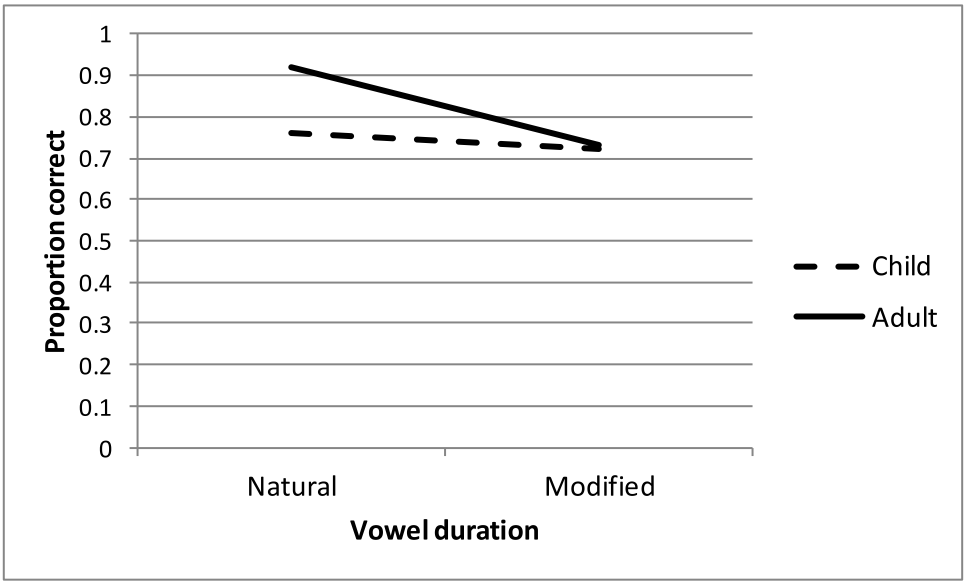


Figure 4.8: Accuracy scores for Greek speakers of both age groups on the AB-X Task.

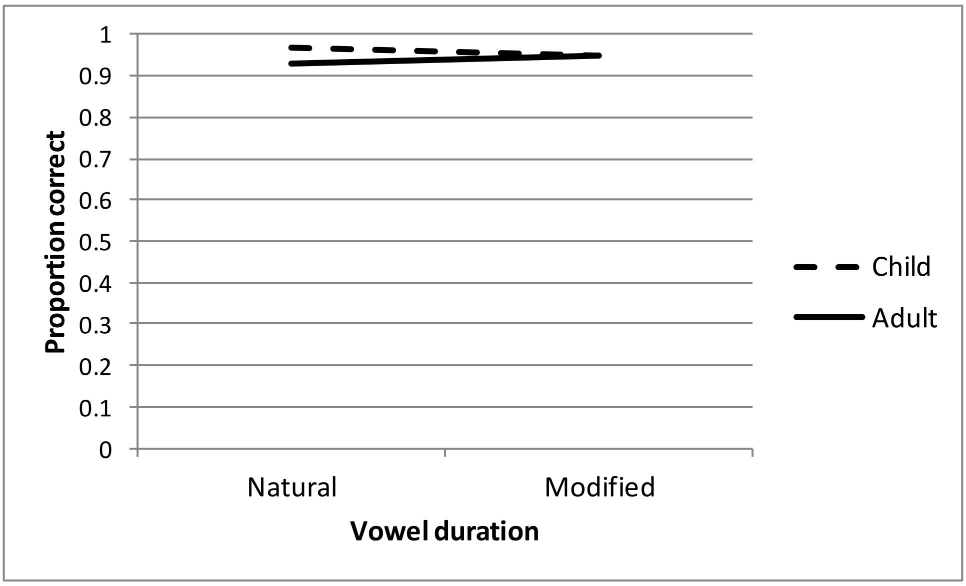


Figure 4.9: Accuracy scores for English speakers of both age groups on the AB-X Task.

4.2.2 Post-training results

For the post-training results, correct responses were averaged across participants for the two training groups, age groups and task conditions. Analyses included the pre-training and post-training test accuracy scores for the two Greek groups (child and adult) only. Accuracy scores were analysed with a mixed analysis of variance (ANOVA) with Age and Vowel Length (duration) as between factors and Training phase (pre- and post-training) as within factor.

Following the use of a 2 week high-variability phonetic training program, performance improved for both Greek adult and child groups as revealed by post training tests. However the effects were most pronounced for the child Greek speaker group. Post-training test results show improvement compared with pre-training scores. Child groups performed better overall than the adult groups. Stimulus type and task type, however, seemed also to affect performance across participant groups.

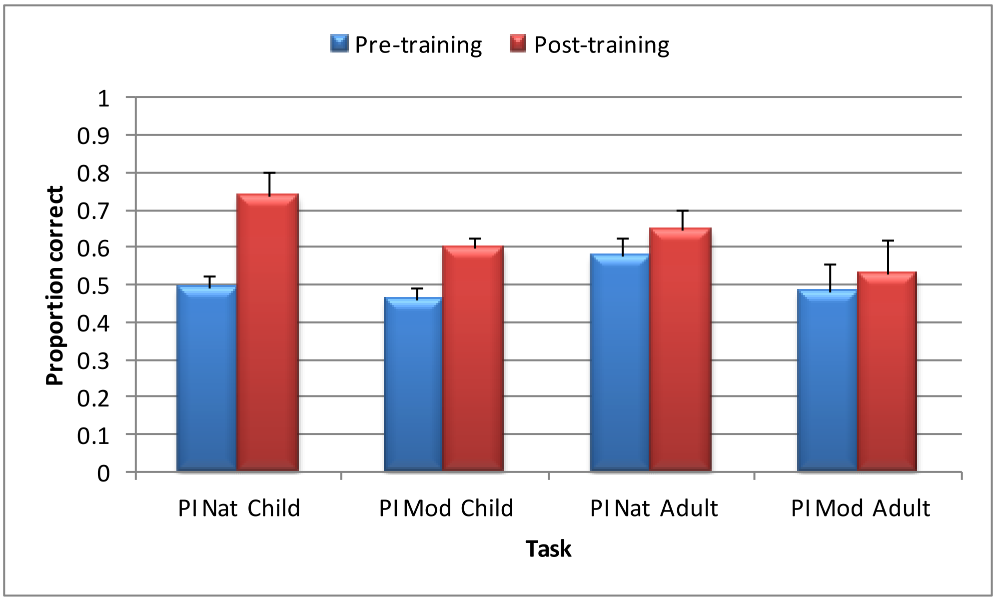


Figure 4.10: Training results for child and adult participants showing pre- and post-training effects for the Perceptual Identification (PI) task and for Natural (Nat) and Modified (Mod) duration conditions.

Table 4.2: Mean accuracy scores and SD for child and adult participants for Perceptual Identification (PI) and Auditory Discrimination (ABX) tasks for Natural (Nat) and Modified (Mod) duration condition, pre and post training.

Group		PI Nat n=45		PI Mod n=45		ABX Nat n=45		ABX Mod n=45	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek child (n=20)	Pre-training	0.49	0.06	0.46	0.05	0.76	0.10	0.72	0.11
	Post-training	0.74	0.05	0.60	0.03	0.90	0.05	0.89	0.04
Greek adult (n=20)	Pre-training	0.58	0.07	0.48	0.11	0.92	0.03	0.73	0.05
	Post-training	0.65	0.05	0.53	0.09	0.94	0.05	0.91	0.05
Mean		0.61	0.06	0.52	0.07	0.88	0.06	0.81	0.06

Specifically, both Greek adult and child groups revealed significant improvement following the training intervention for the perceptual identification task ($F_{(1,36)}$= 47.213, p<.001, see Figure 4.10 and Table 4.2). A significant main effect of Age showed that Greek children had higher improvement compared to Greek adults ($F_{(1,36)}$=10.170, p<.05, see Figure 4.10 and Table 4.2). A main effect of Vowel Length revealed that perceptual identification of natural vowel duration condition was better than for modified vowel duration condition ($F_{(1,36)}$=562.666, p<.001, see Figure 4.10 and Table 4.2). The interaction between Training and Training Group (i.e. trained with natural or modified duration stimuli) was statistically significant since those participants who had received training using modified duration stimuli showed better performance post-training compared with those who had been trained using the natural duration stimuli ($F_{(1,36)}$=17.954, p<.001, see Figure 4.14, 4.15, 4.16, and 4.17). The three-way interaction was statistically significant ($F_{(1,36)}$=21.093, p<.001), indicating that Greek adults showed less of a difference in post-training improvement in both natural and modified duration conditions compared to pre-training. Greek children, however, showed a larger post-training improvement in both natural and modified duration conditions with a more marked difference for the natural condition compared with the modified duration condition (see Figures 4.11 and 4.12 below).

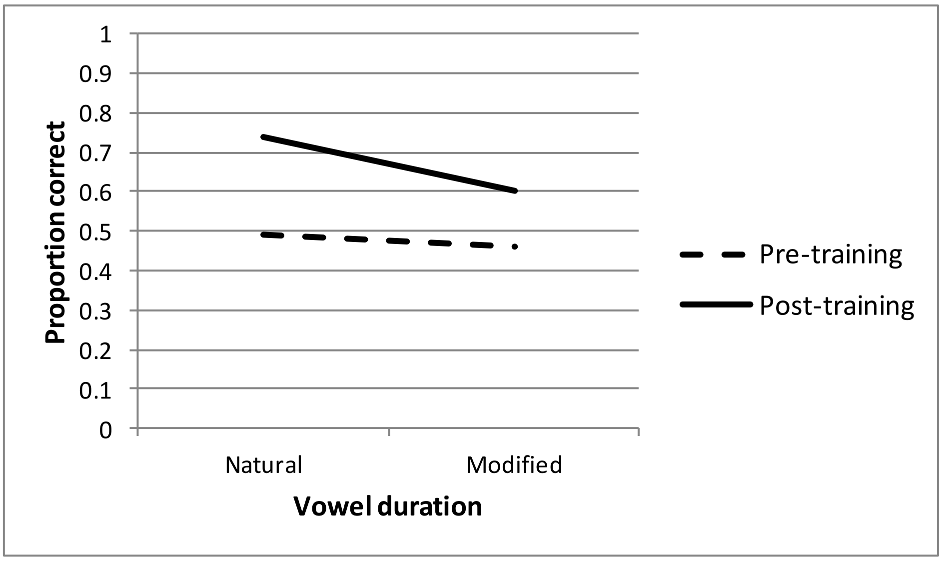


Figure 4.11: Accuracy scores for Greek child speakers for pre- and post-training on the Perceptual Identification Task.

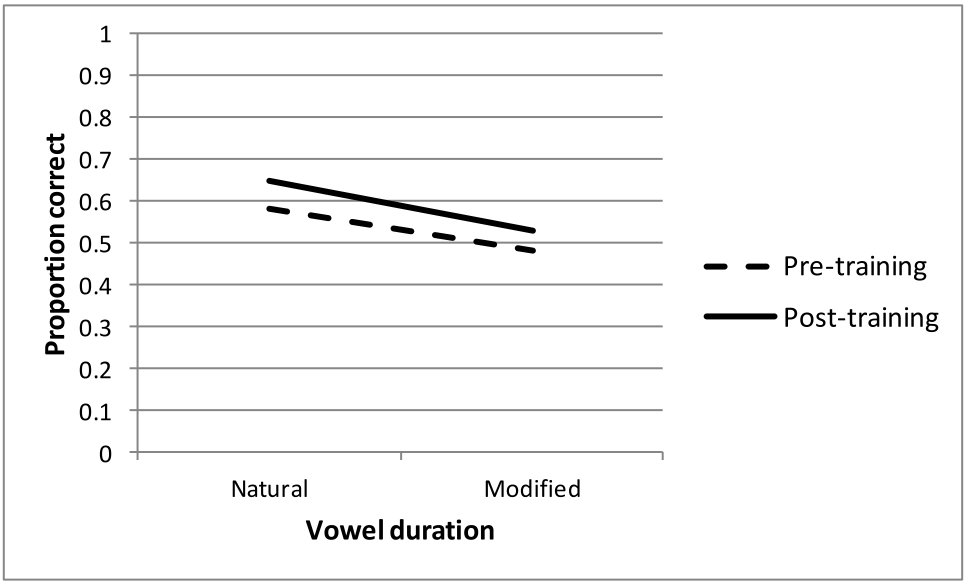


Figure 4.12: Accuracy scores for Greek adult speakers for pre- and post-training on the Perceptual Identification Task.

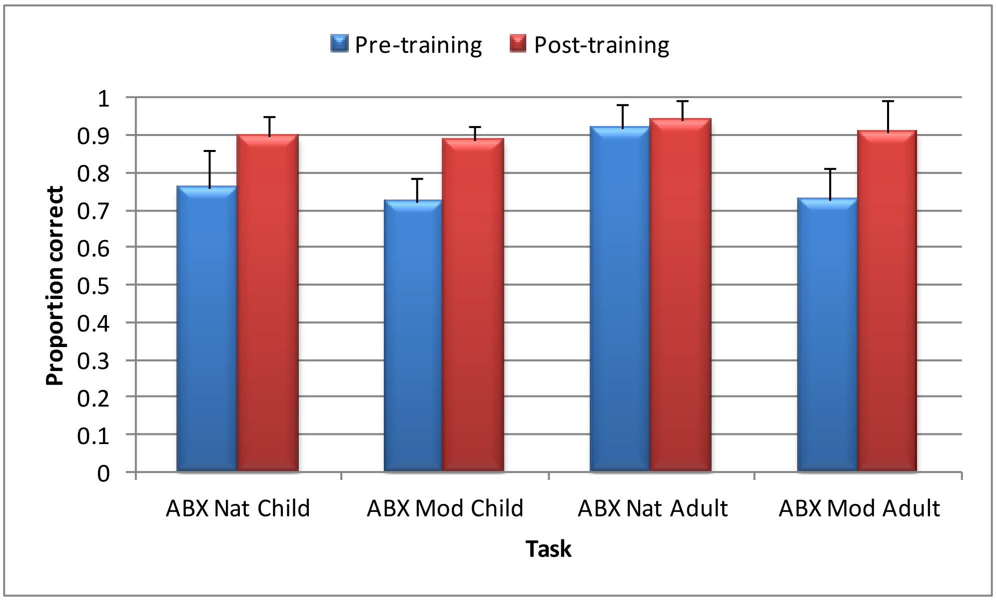


Figure 4.13: Training results for child and adult participants showing pre- and post-training effects for the Auditory Discrimination (ABX) task and for Natural (Nat) and Modified (Mod) duration conditions.

Performance also significantly improved for the AB-X discrimination task as a result of training ($F_{(1,36)}$=52.906, p<.001, see Figure 4.13 and Table 4.2). Discrimination of natural vowel duration words was better than for modified vowel duration words ($F_{(1,36)}$=169.889, p<.001, see Figure 4.13 and Table 4.2), which also mirrors the pattern found in the perceptual identification task results. Both groups improved, but child participants showed the most pronounced improvement in the post training tests as reflected in a significant Training x Age interaction ($F_{(1,36)}$=8.268, p<.05, see Figure 4.13 and Table 4.2). Training improved discrimination performance for both stimulus types (natural and modified) however, those who were trained using the modified stimuli showed a greater improvement compared to those who had been trained using the natural stimuli. Therefore, the Training x Vowel Length interaction was significant ($F_{(1,36)}$=4.615, p<.05, see Figure 4.13 and Table 4.2). Figures 4.14, 4.15, 4.16 and 4.17 display the above results for child and adult groups based on the training condition applied to each group. No other main effects or interactions were significant.

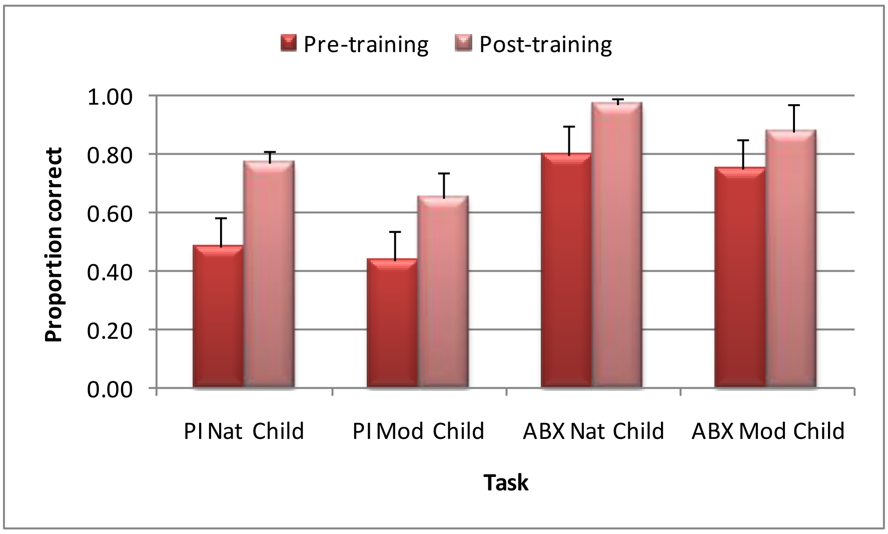


Figure 4.14: Child group trained with Natural Stimuli (n=10 for each group).

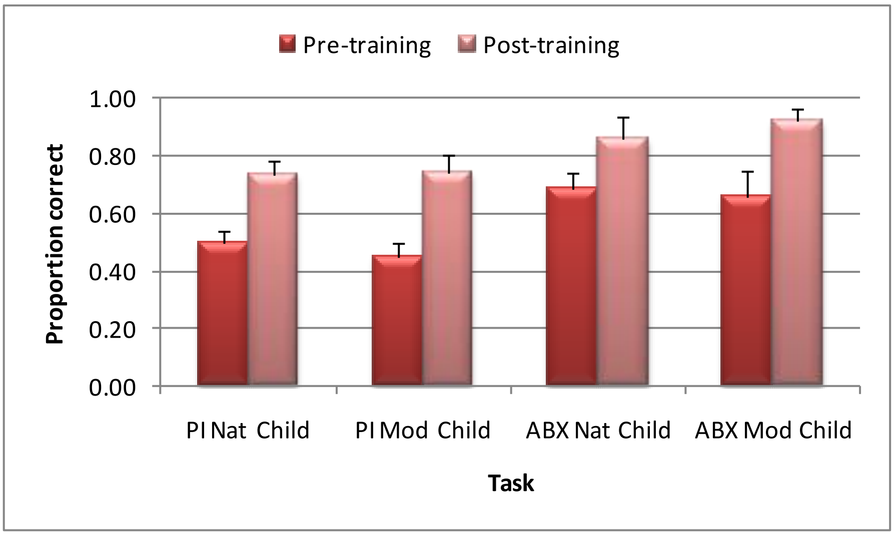


Figure 4.15: Child group trained with Modified Stimuli (n=10 for each group).

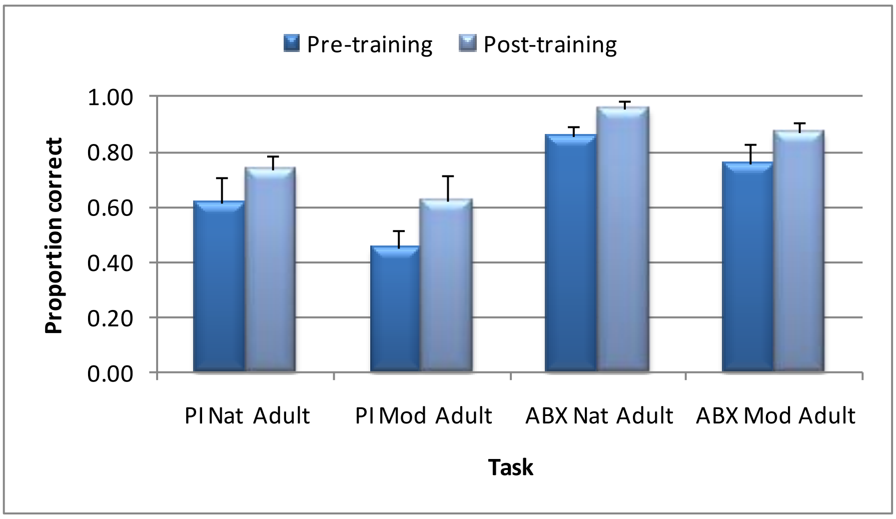


Figure 4.16: Adult group trained with Natural Stimuli (n=10 for each group).

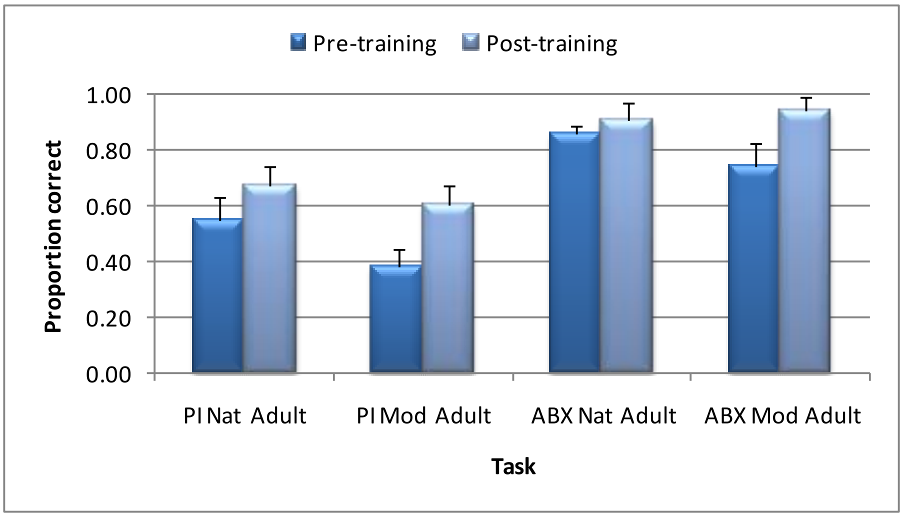


Figure 4.17: Adult group trained with Modified Stimuli (n=10 for each group).

In order to account for the fact that different participant groups had different starting points (as revealed by the pre-training test), difference scores were also calculated; i.e. post-training minus pre-training scores (resulting in % improvement scores). Figures 4.18 and 4.19 show the percent improvement for the child and adult groups.

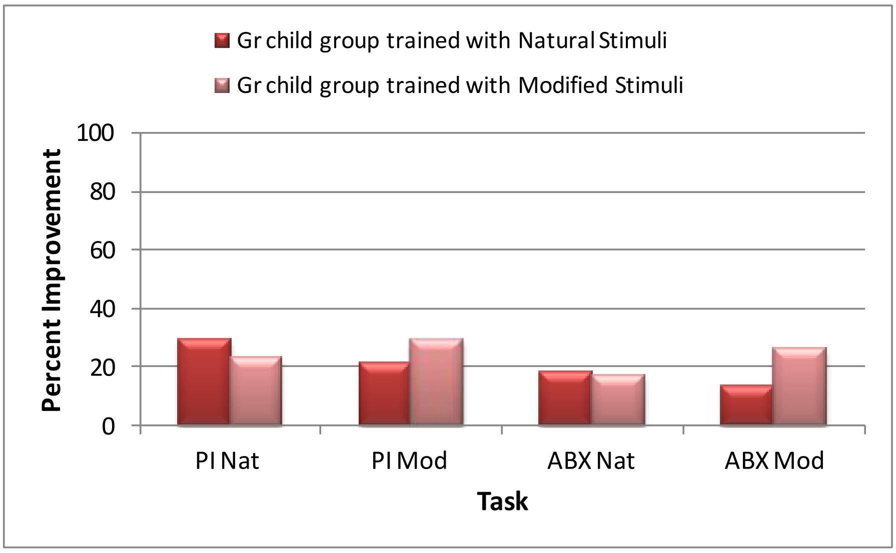


Figure 4.18: Difference scores between pre- and post-training tests for child participants.

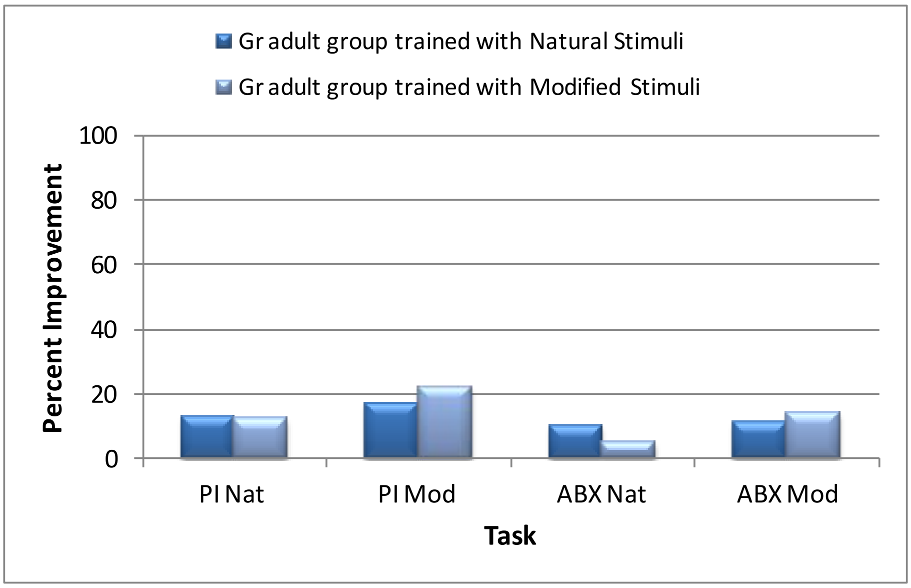


Figure 4.19: Difference scores between pre- and post-training tests for adult participants.

A repeat of the analysis using proportion improvement (i.e. post-training minus pre-training scores) yielded the same results as revealed by post-training statistical analyses.

4.2.3 Training scores and individual differences

A number of recent studies have highlighted the existence of individual differences, for perception of speech sounds, even following training intervention (e.g. Lengeris & Hazan, 2010; Hazan *et al*., 2005, Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba, & Chung, 2006; Golestani & Zatorre, 2009; Lee, Perrachione, Dees, & Wong, 2007; Bradlow *et al*., 1999). Although not the focus of this study, this section presents scores from the individual training sessions for each participant (child and adult) for all training sessions. It is evident that performance improves between Session 1 and Session 10 of the training sessions for Greek child and adult groups alike. Figures 4.20, 4.21, 4.22, and 4.23 below show the learning trend for each participant and experimental condition.

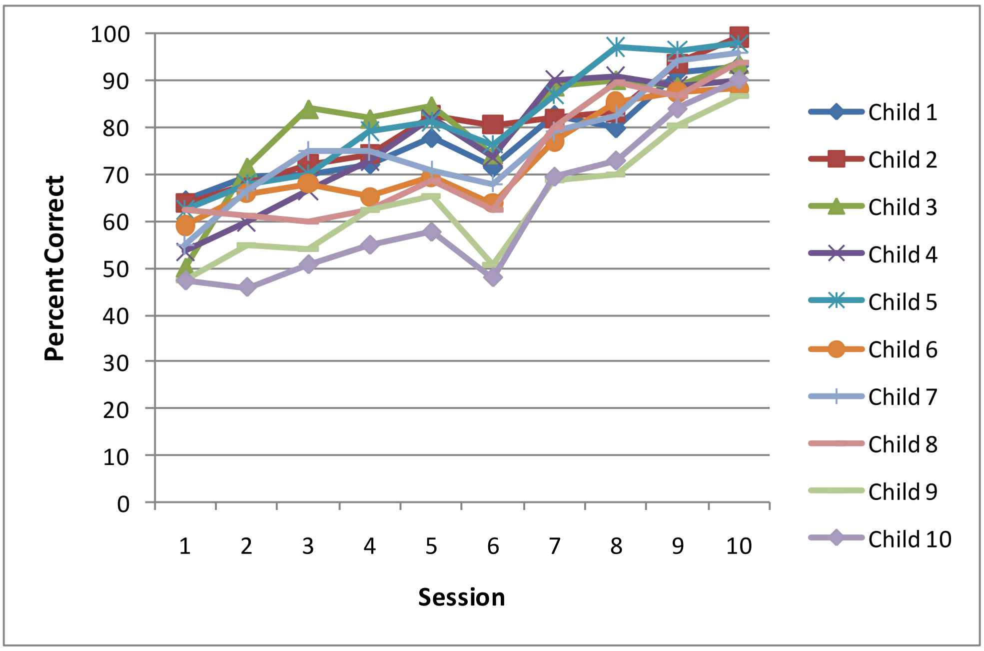


Figure 4.20: Child group trained with Natural Stimuli.

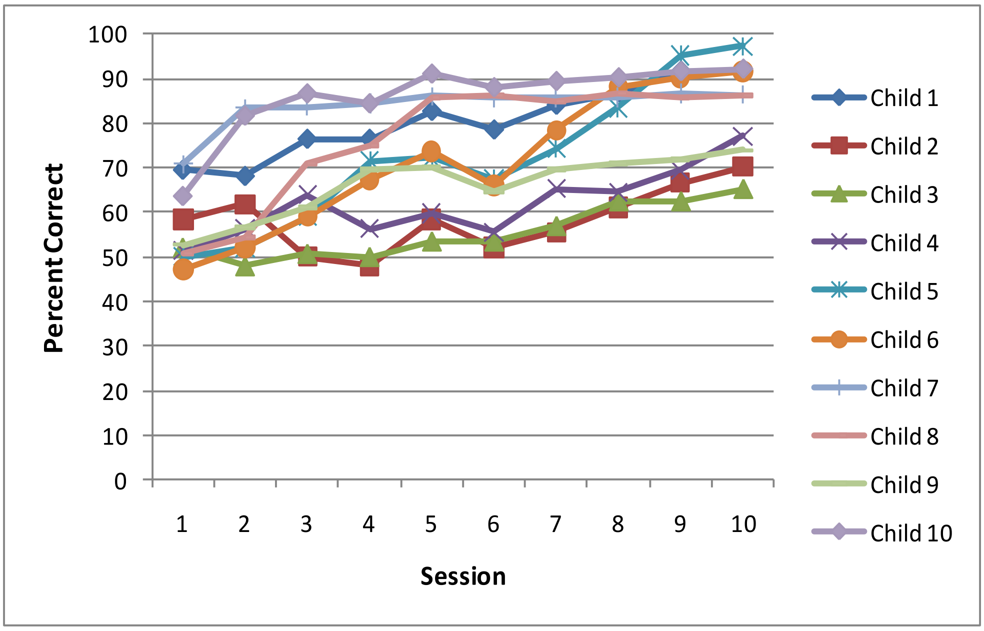


Figure 4.21: Child group trained with Modified Stimuli.

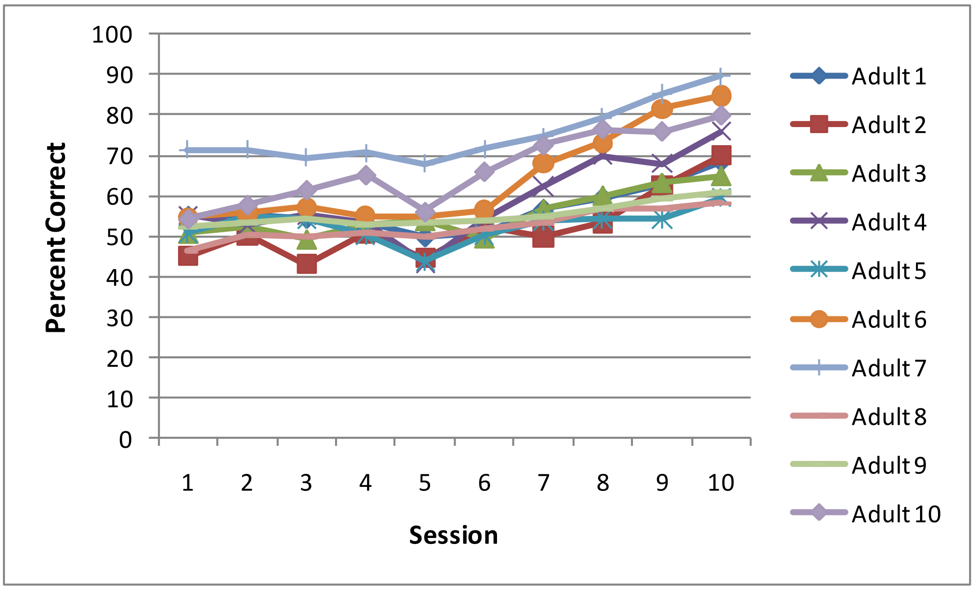


Figure 4.22: Adult group trained with Natural Stimuli.

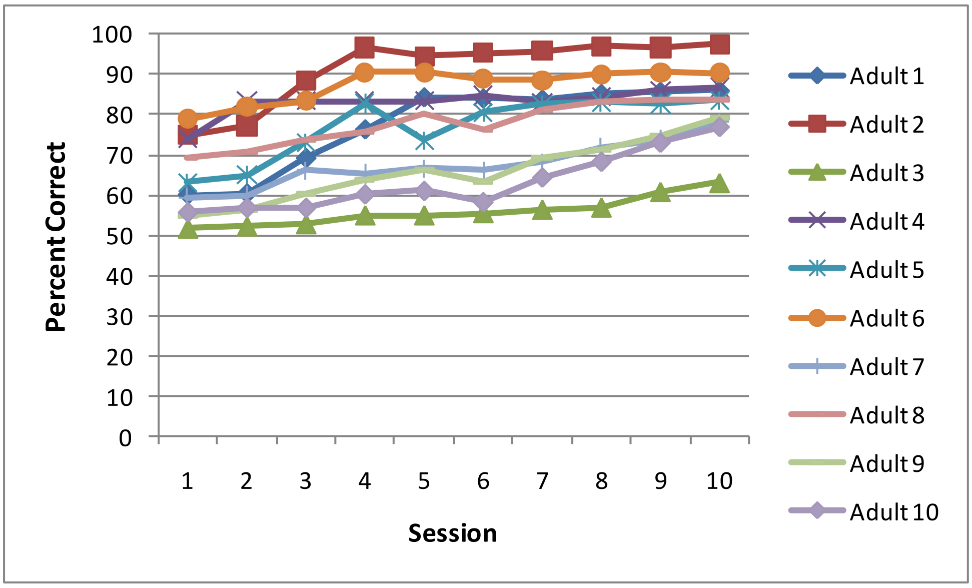


Figure 4.23: Adult group trained with Modified Stimuli.

A general trend is a learning curve occurring in the middle of the training programme where participants (child participants more than adult participants) appear to perform worse for session 5 or 6 than earlier sessions as performance drops (a 'dip' in the curve, particularly obvious for the child participants, see Figure 4.20 and 4.21), but improvement is evident in the following sessions. One plausible explanation for this 'dip' in performance could be that day 6 is after the weekend break between the first and the second week of the training period. An additional observation is how consistently well all child participants trained with natural stimuli are doing on the final day (see Figure 4.20). In contrast, the adults are much more variable, even at the end (see Figure 4.22). This variability is in line with relevant theories about adult L2 learners, for example, Flege *et al*. (2003)'s SLM which emphasizes age effects. This data, however, highlight that although at group level adult participants appear to have greater difficulty with L2 vowel perception (as shown with identification and discrimination tasks in pre- and post-training results), at individual level it is clear that some individuals do perform better than others. In other words, training can have slightly different effects on some individuals which also confirm previous relevant suggestions in the literature (e.g. Lengeris & Hazan, 2010; Hattori & Iverson, 2009; Iverson *et al*., 2008; Iverson & Evans, 2009;

Hazan *et al*., 2005, 2006; Golestani & Zatorre, 2009; Lee *et al*., 2007; Bradlow *et al*., 1999; also Diaz, Baus, Escera, Costa, & Sebastian-Galles (2008) on learners' differences at anatomical and functional level). Also, although child participants evidently start from lower accuracy levels, improvement is higher compared with adult participants. These graphs mirror the pre- and post-training test results discussed earlier.

4.3 Discussion

The primary goal of this study was to examine the identification and discrimination of L2 cues that are not phonemically relevant in the L1 and investigate whether targeted perceptual training could help to improve L2 cue weighting. Specifically, Greek speakers of L2 English were tested on cue weighting for the English /iː/ and /ɪ/ (tense-lax) distinction. Two age groups were considered in order to identify any maturation constraints for either group, especially following the training intervention. Five main points can be identified regarding the identification and discrimination of the English tense-lax distinction by Greek speakers of L2 English and how cues are weighted as well as how cue weighting can be re-learnt.

Firstly, as also discussed in previous chapters, L2 cues are weighted differently between Greek speakers of L2 English and native English speakers. Greek L2 speakers show preference in certain phonetic cues (duration vs. spectral) confirming perceptual interference between L1 and L2 phoneme categories. This limited access to all L2 cues complements the assumption that certain L2 categories assimilate into L1 categories (Best, 1995a). Although spectral cues are not normally accessed by Greek speakers of L2 English, they can be perceptually detected (by adult and child speakers) when forced through task type.

As shown in Chapter 2 and confirmed with the pre-training tests in this study, although English native speakers use spectral cues as primary and duration cues as secondary cues, Greek speakers of L2 English use duration as primary cues while performance is impaired for those tasks where identification and discrimination is based solely on spectral cues. The spectral vowel qualities that are present do not seem perceptually detected with identification task results at chance level for the Greek learners of L2 English groups, compared with ceiling effects by the native English groups. This indicates that cues are weighted differently between the L1 and L2 English groups, but also native language cues (e.g. primary cues in Greek as L1) do not always predict cue weighting in the L2. Instead, cues that are not phonemically relevant in the L1 may be perceptually identified and treated as primary cues in L2 vowel quality identification and discrimination. This is true for both Greek adult and child groups. Interestingly, although duration is not phonemically relevant in Greek vowel discrimination, it appears that the Greek speakers still used duration cues to make the vowel distinction in English.

Second, high variability perceptual identification training can alter perceptual cue weighting. Specifically, targeted training where certain cues are removed (e.g. duration cues) and participants are forced to attend to relevant (e.g. spectral) cues appears to be a highly successful approach to L2 learning. The perceptual training procedure reveals improvement across all tasks and for both adult and child groups. It appears that the combination of multi-talkers used for the stimuli presentation as well as the minimal pair arrangement that involved an identification task type, form a robust method for perceptual identification and learning. Results also confirm that for a trained speech sound contrast it is possible for L2 learners to generalize even to untrained words. This method has been used effectively for Japanese L2 speakers of English (e.g. Magnuson, Yamada, Tohkura, Pisoni, Lively, & Bradlow, 1995; Uther, Uther, Athanasopoulos, Singh, & Akahane-Yamada, 2007) and Finnish (Ylinen *et al*., 2010) and from this data, it also seems to help Greek speakers of L2 English. Thus, high variability can be considered a reliable method for perceptual training. This data could be interpreted in light of the 'attention to dimension' (A2D) models of speech perception (e.g. Iverson & Kuhl, 1995, 1996; Goldstone, 1993, 1994; McClelland, 2001; Nosofsky, 1986), which emphasize the reallocation of the L2 learner's attention on the 'relevant' dimension. Perceptual training could therefore be claimed to shape the L2 learner's perceptual space and shift attention toward the relevant cues (e.g. Francis & Nusbaum, 2002; Kuhl & Iverson, 1995; Jusczyk, 1989).

Third, a significant outcome is the quality of the training stimuli used. The stimulus types used throughout the training sessions (natural versus modified duration stimuli) appear to affect post training performance. Performance was improved across all tasks for those groups trained with stimuli of modified duration; however, improvement was not as pronounced in the case of tasks that involved modified duration stimuli for the groups trained with only natural duration stimuli. Therefore, in the condition where training used natural vowel stimuli (i.e. involved both duration and spectral cues) it appears that Greek L2 speakers of English were likely to weight duration cues as primary and spectral cues as secondary. In the condition where training used modified duration stimuli (i.e. involved spectral cues only) performance was not only improved for the modified duration stimuli tasks but also for the tasks that made use of natural duration stimuli in the post-training test. Therefore, cue weighting can be re-learnt, cues previously perceived as secondary can be rearranged into primary cues, however it appears that the stimuli type used in the training is of high importance for this cue weighting perceptual rearrangement during the training intervention. Future experiments including a larger number of participants for the various experimental conditions would be necessary in order to further explore this assumption.

Fourth, comparison of child versus adult groups shows age differences in training intervention, pointing out the maturation constraints involved. Improved post training results for the adult group demonstrate a phonetic capacity for adequate perceptual adaptation and sufficient neuroplasticity for such adaptation. Therefore, high variability training can improve perceptual identification and discrimination for phonetic cues in adult groups. Although training significantly improved performance for both groups, the degree of improvement was significantly larger for the child than the adult group, even despite differences in years of English language education and at very obviously different starting points. This outcome could be discussed in relation to earlier assumptions (see Chapter 2 and 3) that age effects between adult and child L2 learners' task performance could be interacting with the difference in proficiency levels between the two age groups. The current study, however, seems to challenge earlier assumptions, to some extent, given the children's marked post-training improvement. Results here could suggest that proficiency level cannot be considered a restricting factor for the degree of improvement following perceptual training. It is likely that plasticity rather than proficiency effects should be considered as contributing to such learning outcomes, given the children's lower baseline levels.

Finally, individual differences evidently exist. However, there are clear learning trends during the course of the training programme, reflected in the pre- and post-training results. The child group demonstrates a more robust capacity for perceptually rearranging their phonetic cue weighting. This would suggest children are more easily susceptible to inducing plastic changes in their phonetic categories. The results are also in support of the speech learning models that suggest a consolidation of neural and perceptual phonetic categories with increasing age (Kuhl *et al.*, 2008; Flege, 2002).

The general assumptions as revealed in this study suggest that L2 learners often weight cues differently compared to L1 speakers. This perceptual cue weighting can be modified through high variability perceptual training. Adult and child groups appear to have access to perceptual cue weighting modifications, although neuroplasticity for child groups seems to be more malleable.

Chapter 5

Phonological processing in English by Greek learners of L2 English: evidence from tasks involving reading.

5.1 Introduction

Reading researchers and educators have identified the utility of reading and its usefulness for language learners (e.g. Bernhardt, 2005; Kucer, 2005; Smith, 2004) as well as the role of reading for L2 learners in acquiring and developing their L2 language skills (e.g. Krashen, 1993). Reading is one of the main skills developed while learning a language (L1 or L2 alike), among listening, speaking, and writing skills. Speech perception and reading differ considerably: written words can be viewed as a whole but spoken words are spread over time, written words have clear boundaries but it is in general harder to identify the boundaries of spoken words, or even spoken words taken out of context can be harder to identify (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) compared to written words.

Speech offers cues (e.g. prosodic cues) that are different to those offered in the case of written text. In English for example, each of the two words in the minimal pair '*seat*' vs. '*sit*' is pronounced differently (more specifically, each word involves a different vowel sound, /iː/ and /ɪ/ respectively, and they involve cues, e.g. spectral and duration cues, that differentiate one word from the other in speech as discussed in previous Chapters), while the two words are spelt differently (i.e. the vowel in each minimal pair counterpart is orthographically represented as '*ea*' and 'i' respectively). Other examples, however, include pairs that also have different orthographic representation but are pronounced exactly the same, such as 'blew' vs. 'blue', and are called *homophones* (i.e. they sound the same, also called *heterographs* in the literature referring to their one pronunciation but two spellings). An example of homophones in Greek could be the words '*σήκω*' (/siko/, stand up) versus '*σύκο*' (/siko/, fig) which are both pronounced exactly the same, /siko/, but are spelt differently *(σήκω* vs*. σύκο*). Another category is the *homographs* (i.e. same spelling but different pronunciation based on context), for example the word 'wind' (pronounced as either /wɪnd/ or /waɪnd/). For the L2 learner, developing their reading skills in two different languages (L1 and L2) is thus a considerable task and the above examples only briefly demonstrate the implications involved between the orthographic systems of different languages and how they relate with the phonology of those languages.

Landerl, Wimmer, & Frith (1997) examined reading and phonological processing skills by English and German children. Although the main target group was dyslexic children compared with unimpaired child groups, their findings point to language-related implications than group-related ones. Specifically, Landerl *et al*. (1997) give evidence that it is easier for children to learn the relevant grapheme-phoneme rules underlying reading in cases of transparent languages (e.g. German, or even Italian, Spanish, Greek (for a review see Goswami, Ziegler, & Richardson, 2005), where grapheme-to-phoneme conversion mappings are more consistent), compared to developing reading skills in more opaque languages (with less consistent grapheme-to-phoneme conversion mappings, e.g. English). They argue that phonological decoding may not be a demanding task as such but it is the transparency level of the target language that may affect reading processes.

A number of previous studies have highlighted that phonological processing skills facilitate reading processing (e.g. Adams, 1990; Badian, 1998; Bryant, 1986; de Jong & van der Leij, 2002; Denton, Hasbrouk, Weaver, & Riccio, 2000). Hulme, Hatcher, Nation, Brown, Adams, & Stuart (2002) suggested that phonemic awareness can be a reliable predictor of word reading ability. However, although phonological processing skills are indeed imperative for reading processing, further research has pointed to the fact that it is both orthographic processing[34] and phonological processing that are mutually facilitated. Ziegler, Stone & Jacobs (1997) conducted a statistical analysis to determine reading consistency in 2,694 monosyllabic English words. They calculated reading consistency, i.e. graphemes that can be pronounced in multiple ways, and spelling consistency, i.e. phonemes that can be spelled in multiple ways. From the words used in the analyses, they found that 72.3% of the words were inconsistent on the phoneme level and only 30.7% of the words were inconsistent on the grapheme level. These findings highlight the complexity of opaque orthographies, such as English, when it comes to reading.

Considering the Landerl *et al*. (1997) findings mentioned earlier in view of the Ziegler *et al*. (1997) results, it is possible to hypothesize that transparency levels of a language could therefore affect reading processes of L2 learners, based on the orthographic (transparency) relation between the L1 and L2. L2 learners of an opaque language (e.g. English) whose L1 background is a transparent language (e.g. Italian, Spanish, Greek) could be expected to have a higher degree of reading difficulty (revealed by higher error scores or longer RTs) as compared to a reverse L1-L2 language relation (i.e. L1 being an opaque and L2 being a transparent language).

It has been reported that learning to read "even affects the purely auditory perception of speech by introducing subtle influences related to spelling" (Pattamadilok, Knierim, Kawabata-Duncan, & Devlin, 2010, p. 8435). For example, Experiment 2 (discussed in Chapter 3) showed that for the perceptual identification (PI) task that involved pictures, Greek participants (adult and child groups alike) performed worse than in the same task-type which involved orthographic word stimuli compared with the respective English control groups. This finding suggests that the way written words are spelt provide

[34] '*Orthographic processing'* is defined as awareness of permissible letter patterns in a language (e.g. Perfetti, 1984).

orthographic cues[35] that could affect speech perception outcomes. For example, Seidenberg & Tanenhaus (1979) found that two similarly spelt words (e.g. *hear – near*) were recognized faster by participants compared to dissimilarly spelt words (e.g. *hall - bawl*) (see also, Tanenhaus, Flanigan, & Seidenberg, 1980; Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981). The Seidenberg & Tanenhaus (1979) findings could therefore lead to the assumption that orthographic similarity or dissimilarity could be factors that could affect L2 learners' reading performance. Orthographic similarity, however, does not guarantee that two words will rhyme in an orthography that is less transparent. For example, in English the word pair '*weight – height*' employs orthographic similarity but in this case such similarity can be misleading as the two words do not rhyme.

The orthographic relationship between a L1 and a L2 could not only be an indicator of reading processing and performance but could actually give valuable insights for the existence of orthographic cues that could either aid or inhibit L2 successful reading. This view stems from two studies which investigated the role of phonological and orthographic processes in Chinese (L1) – English (L2) (Wang, Perfetti, & Liu, 2005) and Korean (L1) – English (L2) (Wang, Park, & Lee, 2006) children. Participants in both studies were given reading tasks in both languages (Chinese – English and Korean – English, respectively), and aimed to examine phonological and orthographic processing skills when reading. Results from the Wang *et al*. (2005) study showed that although Chinese (L1) tone significantly affected English (L2) non-word reading, orthographic processing did not predict L2 reading. Similarly, results from the Wang *et al*. (2006) study also supported the view that L1 phonological processing transfers in the L2 but L1 orthographic processing does not significantly contribute to L2 word reading. The authors' explanation in both studies for this lack of orthographic transfer focuses in the

[35] The term '*orthographic cues*' here refers to cues /indications provided by written text (strings of letters or graphemes) in relation to phonemic representations and text-to-sound mappings in a given language. This term has also been used in the literature to define different orthographic manipulations in various experimental conditions (see e.g. Snowling & Frith, 1981), but such other definitions do not necessarily apply to how it is defined within this context.

existing visual difference in orthographic representations between the respective L1 and L2. Specifically, Chinese and Korean orthography consists of nonlinear, square-like characters as opposed to English orthography which is a linear, alphabetic system. Therefore, the fact that speakers of L1 Chinese or Korean have not learnt mappings between phonemes and alphabetic-graphemes during L1 reading acquisition, leads to this lack of orthographic transfer observed by Wang *et al*. (2005, 2006).

Consequently, a reverse assumption could be that L2 learners of an alphabetic language whose L1 is also an alphabetic language could be expected to show orthographic transfer effects. This could be in the form of orthographic cues. Such orthographic cues, however, could either aid or inhibit correct L2 mappings. That is, speakers of an L1 may have learnt one-to-one grapheme-phoneme mappings while the L2 may employ one-to-many grapheme-phoneme correspondences. For example, in Greek the grapheme 'α' is always pronounced /a/ while in English the same grapheme 'a' can have a variety of pronunciations, e.g. /a/ as in 'cat', /ɔ/ or /ɒ/ as in 'water' (based on dialect), /ə/ as in 'natur<u>a</u>l (natʃrəl)', /eɪ/ as in 'bake' or 'face', or even be silent (as in 'bread'). This one-to-one (L1) versus one-to-many (L2) grapheme-phoneme correspondence is therefore very likely to cause confusions at reading performance for the L2 English learner.

Experiments presented in previous chapters showed results of L2 speech sounds being assimilated or mapped on L1 speech sounds, and auditory cues being weighted differently between L1 and L2 English speakers. Experiment 2 in particular, emphasized the possibility of orthographic cues being embedded in the orthographic representations of words that could 'aid' auditory speech perception. Previous experiments of this thesis considered aspects of perceptual assimilation demonstrating mappings of L2 speech sounds onto existing L1 sounds; for example, when two distinct L2 categories are represented by a single L1 category and therefore the two L2 categories are assimilated into the one L1 category (e.g. Brown, 2000; Kuhl, 1993, 2000; Best, 1995, 2001; Flege, 1987; 1995). The fact that the L2 learner is also a L2 reader could postulate that similar mappings may also occur on the orthographic level. When the same letters (graphemes) represent similar (but not the same) speech sounds between the L1 and the L2, in other words there is a ‘shared orthography’ between two languages, and that could result in confusions for the L2 learner (e.g. Oller & Ziahosseiny, 1970). When the same grapheme is shared between two languages but representing different speech sounds then in order to overcome any perceptual confusion, the L2 learners will need to disassociate the grapheme-phoneme accosiations that already exist in their L1 (e.g. Jared & Szucs, 2002; Treiman & Cassar, 1997).

On the word level, there is support that consistency of mappings between orthography and phonology influence visual word identification (e.g. Gottlob, Goldinger, Stone & Van Orden, 1999; Hino, Lupker & Pexman, 2002; Stone, Vanhoy & Van Orden, 1997). In the monolingual domain for example, it has been found that performance is inhibited in the case of words where orthography can map onto multiple phonological codes (e.g. ‘*read*’ pronounced as [rid] or [rɛd], and ‘*lead*’ pronounced as [lid] or [lɛd]) (see e.g. Gottlob *et al*., 1999; Hino *et al*., 2002; Stone *et al*., 1997). A similar outcome has also been supported when phonological codes can map unto multiple orthographies which is the case of homophone words (e.g. [meid] for ‘*maid*’ or ‘*made*’) (Pexman, Lupker & Jared, 2001; Pexman, Lupker & Reggin, 2002; Stone *et al*., 1997). This demonstrates that phonological processing is influenced by orthographic cues and this becomes evident at reading task performance.

In the bilingual domain, cross-language activation has been investigated in studies that employ word naming tasks and in order to produce an appropriate response, the full phonological code is required (e.g. Jared & Kroll, 2001; Jared & Szucs, 2002). For example, Jared & Szucs (2002) tested English – French bilinguals on interlingual homographs (i.e. words that share form but not meaning across different languages, e.g. ‘*pain*’ meaning ‘bread’ in French) and observed cross-language phonological interference, supporting the claim that bilinguals activate phonological codes from both languages simultaneously, even in the case of the target language being their more dominant one. Similar observations from other studies have also been reported, particularly examining the relation between orthography, phonology and semantics (e.g. Altenberg &

Cairns, 1983; De Groot, Delmar & Lupker, 2000; Dijkstra, De Bruijn, Schriefers & Ten Brinke, 2000; Van Heuven, Dijkstra & Grainger, 1998). Word naming has generally been observed to be affected by orthography in the case of orthographically identical words (e.g. interlingual homographs) (e.g. Beauvillain & Grainger, 1987; Dijkstra *et al*., 1998) and also in the case of cognates (i.e. words that share form and meaning across different languages) (e.g. Dijkstra *et al*., 1999; Lemhöfer, Dijkstra & Michel, 2004).

It has been suggested that when children start to read they cannot adequately separate speech sound and spelling (e.g. Treiman & Cassar, 1997; Jared & Kroll, 2001; Jared & Szucs, 2002). L1 phonological representations are suggested to be automatically activated in reading (e.g. Jared & Szucs, 2002) which leads to an orthographic and phonological co-activation. This makes readers more 'susceptible to unwanted interference' (Landerl, Frith, & Wimmer, 1996, p.12). A study by Schwartz, Kroll & Diaz (2007) has demonstrated that the consistency of orthographic and phonological mappings between two languages (in this case, English and Spanish) influence reading processing for the L2 learner. Additional research has shown that orthographic information can override phonetic information which can lead L2 learners to ignore or misperceive acoustic information (e.g. Erdener & Burnham, 2005; Hallé, Chéreau & Segui, 2000). Thus, if orthographic information can influence the processing of phonological information, in the case of L1 – L2 shared orthography (or shared symbols), this could lead L2 learners to misperceive L2 sounds and concequently misread L2 words.

Considering the orthographic systems between Greek and English, it is evident that some symbols in the two systems are identical (i.e. shared orthography) and others are not. However, some of the symbols that are identical in orthographic form do not correspond to the same phoneme sound. This observation can lead to the prediction that orthographically similar word pairs in English that do not rhyme (e.g. *weight [weit] – height [hait]*) could cause interferences since 'e' in Greek can only be pronounced as [e], not as [a], unlike in English (see Appendix 4b for a list of the letters of the Greek alphabet and the corresponding English letters and letter sounds).

The shared orthography between Greek and English in relation to the transparency levels for each of the two languages (i.e. Greek being a transparent language and English employing an opaque or less transparent orthography, see Section 1.6.3 for a discussion), could postulate that reading in English can be a rather complicated task for the Greek learner of L2 English. L1 orthographic transfer could be expected based on the existing orthographic similarity, as learning similar components between two languages would cause more difficulty for the L2 learner than learning dissimilar ones (e.g. Oller & Ziahosseiny, 1970; Wode, 1983; Flege, 1987, 1995; Major, 2001). Grapheme-to-phoneme and phoneme-to-grapheme inconsistencies can therefore influence performance of the L2 reader that is manifested by longer reaction times and error scores (e.g. Glushko, 1979; Jared, McRae & Seidenberg, 1990; Lacruz & Folk, 2004; Stone, Vanhoy & van Orden, 1997).

Considering the two reading models discussed in Chapter 1 (section 1.6), namely the '*Dual-Route Cascade model of reading (DRC)*' (Coltheart *et al*., 2001; Coltheart & Rastle, 1994; Rastle & Coltheart, 1998) and the '*Parallel Distributed Connectionist model of reading (PDP)*' (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996; Plaut & Kello, 1999), it is possible to make some hypotheses in the context of rhyming judgment experimental tasks. According to the DRC, words could have faster RTs compared to non-words. Non-words follow the sub-lexical grapheme-to-phoneme conversion (GPC) route but in the context of a rhyming judgment task, for instance, a 'lexical look up' (Rastle & Coltheart, 2000, p. 343) procedure could also be activated as in the case of the non-word 'vint' as opposed to either 'pint' or 'mint'. Non-word rhyming judgment tasks would therefore be expected to produce longer RTs in general. Orthographically similar rhyming pairs (e.g. *plead* vs. *bead*) are also expected to produce faster RTs compared to orthographically dissimilar rhyming pairs (e.g. *hall* vs. *bawl*). The PDP model would make predictions based on the 'weighting' between the units (orthography, phonology and semantics) and the reader's experience. In the case of a familiar word, semantic 'weighting' is also activated, unlike a non-word. RTs could therefore be expected to be longer in the case of non-words. Irregular word spelling (i.e. orthographically dissimilar words) would be expected to trigger longer RTs as semantic 'weighting' is also predicted to take place prior to word reading output.

This chapter[36] considers the cognitive processes of reading words and non-words by Greek learners of L2 English in a series of rhyme and homophone judgment tasks. With a transparent L1 language system as in the case of Greek, Greek learners of L2 English could be slower readers compared to English native speakers. Some alphabetic symbols are shared between Greek and English but not necessarily with sound correspondence in all cases (see Appendix 4b), could lead to the expectation of further confusions for the Greek readers of L2 English. Looking into differences between words and non-words, as well as overall orthographic consistency effects, this study aims to shed light into the implications of reading processing for Greek learners of L2 English.

A main research question is whether orthographic cues are weighted by L2 learners differently to native speakers. Addressing this question, it could be possible to examine whether results in the auditory modality (as discussed in previous chapters) could also be reflected in the visual modality. This study makes the following hypotheses: 1. Differences in orthographic transparency levels between Greek (L1) and English (L2) can cause distortions in reading performance, e.g. for words with similar orthography but different pronunciation (e.g. *weight – height*) where full word reading is required for correct reading performance; 2. Shared orthography effects can also inhibit correct reading performance due to already existing mappings of grapheme-to-phoneme correspondance; 3. Considering the orthographic relationship between Greek (L1) and English (L2), a main hypothesis is that Greek learners could demonstrate a tendency for orthographic transfer through use of L1 orthographic cues at L2 reading.

Previous studies have considered other first languages in relation to L2 English, such as Chinese (e.g. Lee, Huang, Kuo, Tsai, & Ovid Tzeng, 2009),

---

[36] Part of the work comprising this chapter has been presented at the British Association of Cognitive Neuroscience (BACN) Annual Conference, UCL, Institute of Child Health, 2009.

French (e.g. Comeau, Cormier, Grandmaison, & Lacroix, 1999), Italian (e.g. D'Angiulli, Siegel, & Serra, 2001), Korean (e.g. Wang *et al*., 2006), but no other study – to the author's knowledge - has considered reading processes of English words by Greek learners of L2 English in a similar experimental design.

5.2 Method

5.2.1 Participants

Twenty native speakers of Standard Modern Greek aged 20-30 (8 female, 12 male; mean age = 25.4) were recruited from the University of Patras, Greece and twenty native monolingual English native speakers aged 19-28 (15 female, 5 male; mean age = 22.3) were recruited from Brunel University, West London, UK. The English native speakers served as controls. The Greek native speakers had all lived in Greece and had studied English as L2 in school (public and private education; L2 English education mean = 8.7 years). Their level of proficiency in English was advanced in reading and writing (all had received the Cambridge Certificate in Advanced English or Certificate of Proficiency in English) as recorded in the language background questionnaire that all participants completed before testing. None had spent more than 2 weeks in an English-speaking environment. English native speakers received course credit for their participation. Greek learners of L2 English received a small gift of stationery items as reward for their participation and were entered into a prize draw. All participants had normal or corrected-to-normal vision and had no history of learning disability or reading impairment.

5.2.2 Stimuli

***Word Rhyme Judgment*:** 60 word pairs from Howard & Franklin (1996) were used (see Appendix 4a). This task involved 4 conditions: 15 orthographically similar rhyming words (OrthSimRhymWord) (e.g., *plead* - *bead*), 15 orthographically dissimilar rhyming words (OrthDissRhymWord) (e.g., *hall* –

bawl), 15 orthographically similar non rhyming words (OrthSimNonRhymWord) (e.g., *weight – height*), and, 15 orthographically dissimilar non rhyming words (OrthDissNonRhymWord) (e.g., *crane – wine*).

***Non-word Rhyme Judgment*:** 50 non-word pairs from Best (1996), cited in Wydell & Butterworth (1999), were used (see Appendix 4a). This task involved 2 conditions: 25 rhyming non-words (RhymNonWord) (e.g., *yite - pight*), and, 25 non-rhyming non-words (NonRhymNonWord) (e.g., *trosh - desh*). It should be noted that correct pronunciation for non-words is based on English reading rules which allow for one possible correct pronunciation. Therefore, correct pronunciation for rhyming and non-rhyming non-word pairs used in this study is specific and cannot be arbitrary.

Homophone Judgment: 100 homophone word pairs from Coltheart (1980), cited in Wydell & Butterworth (1999), were used (see Appendix 4a). This task involved 4 conditions: 25 Regular Homophones (RegHom) (e.g., *blew - blue*), 25 Irregular Homophones (IrregHom) (e.g., *eye - I*), 25 Regular Non-Homophones (RegNonHom) (e.g., *bound - boned*), and, 25 Irregular Non-Homophones (IrregNonHom) (e.g., *fair - fear*).

5.2.3 Apparatus

Stimuli were presented in a laptop (AMD Sempron) with E-Prime software (Schneider *et al.*, 2002a,b). Visual stimuli were displayed on a 33 x 20 cm monitor. Reaction times (RT) and responses were automatically recorded for each participant through the E-Run[37] software application. Participants responded by pressing relevant keys (‘P’ for a ‘yes’ response and ‘Q’ for a ‘no’ response) on the computer keyboard which triggered the next trial after a 1000 ms delay. A word pair of Courier New font, size 18, black bold type, and white

[37] The E-Run software application is the presentation component of E-Prime 1.1 enabling for millisecond precision of stimulus presentation, as well as synchronizations of stimuli (e.g. visual and/or auditory) and data collection.

background would appear on the computer screen across all experimental tasks and conditions.

5.2.4 Procedure

Participants were tested individually in quiet rooms and in convenient geographical locations. They all completed a language background questionnaire (see Appendix 5 for a sample questionnaire) and gave written consent before participating in the experiment. Participants initially had a short practice session of 10 trials in order to familiarize themselves with the task procedure. After completion of the practice session they started the full experimental tasks. Participants were instructed to read the two words appearing on the screen in each trial and respond as quickly as possible by pressing a relevant key ('P' for a 'yes' response and 'Q' for a 'no' response) on the keyboard.

For the *Word Rhyme Judgment* task participants were asked to judge whether the word pair rhymed (e.g., *plead* - *bead*) or not. For the *Non-word Rhyme Judgment* task participants were asked to judge whether the non-word pair rhymed (e.g., *yite* - *pight*) or not. The participants were also informed that the stimuli were all non-word pairs, but that their task was to try and read them and respond as accurately as possible using the keyboard keys. For the *Homophone Judgment* task participants were asked to judge whether the word pair sounded the same when read (i.e. they were homophones) (e.g., *blew* - *blue*) or not.

Every task started with a set of instructions appearing on the screen which was terminated by the participant through pressing the space bar on the keyboard when ready to start the task. Likewise, at the end of each task, a smile sign appeared and a message notifying the participant that they could take a short break and indicate when ready to start the text task which also included an introductory instructions message at the beginning of the task and concluding screen message at the end. The stimuli within each task were presented in randomised order generated by the E-Run application and the order of task presentation was counterbalanced across the experimental group. Each trial started with the presentation of a fixation “+” for 500ms, followed by a pair of stimulus words which were presented until a response was made (or until 3000ms if no response was made). The inter-stimulus interval was 1,000ms. Altogether the whole experiment took 45 minutes.

5.3 Results

5.3.1 RT Analyses

RTs for the correct responses were averaged across participants for the two language groups (Greek and English) for all tasks and conditions. Where responses deviated more than 2.5 standard deviations (SD) from the mean for each participant, these trials were classed as outliers and the data were excluded from the analysis. Outliers accounted for less than 4% of all the trials. Tables 5.1, 5.2, and 5.3 show the mean reaction times (RTs) in milliseconds (ms) of correct responses and SD, for each task and experimental condition for Greek and English speakers. Data were analyzed using mixed ANOVAs for subject (F_1) and item (F_2) analysis.

5.3.1.1 RT results for Word Rhyme Judgment task

The RT data for correctly selected items in the *Word Rhyme Judgment* task were analysed with a three-way (Group x Rhyme x Orthographic Similarity) analysis of variance (ANOVA). The main effect of Group was significant, $F_{1(1, 38)}$= 34.64, *MSE* = 1730253.482, p<.001; $F_{2\ (1,\ 28)}$= 304.6, p<.001, *MSE* = 140070.27, p<.05, indicating that the English group were significantly faster to respond than the Greek group. The main effect of Rhyme was significant, $F_{1(1, 38)}$= 4.29, *MSE* = 3607125.333, p<.05; $F_{2(1,\ 28)}$= 17.41, p<.001, *MSE* = 149339.71, p<.05, suggesting that rhyming words were easier to identify than non-rhyming words. However, the main effect of Orthographic similarity was not significant ($F_{1(1, 38)}$= .032, *MSE* = 388407.706, p=.86; $F_{2(1, 28)}$= .045, *MSE* = 338072.531, p= .834,) suggesting that the orthographic similarities of words for each trial pair did not affect RTs. The interaction between Group and Orthographic similarity revealed significant results ($F_{1(1, 38)}$= 6.43, *MSE* = 2500310.907, p<.05; $F_{2(1, 28)}$= 4.94, *MSE* = 1671458.25, p<.05). This interaction revealed that the Greek group had longer RTs for the orthographically dissimilar words compared with the orthographically similar words. On the other hand, the English speakers had longer RTs for the orthographically similar words compared with the orthographically dissimilar words (see Table 5.1 and Figure 5.1).

Table 5.1: Mean RTs and SD for Greek and English speakers for the Word Rhyme Judgment Task.

	OrthSimRhym Word		OrthDissRhym Word		OrthSimNonRhym Word		OrthDissNonRhym Word	
	n=15		n=15		n=15		n=15	
Group	Mean(ms)	SD	Mean(ms)	SD	Mean(ms)	SD	Mean(ms)	SD
Greek (n=20)	2719.9	759.2	3811.9	1398.3	3635.4	1163	3078.5	712.7
English (n=20)	1710.1	503.5	1954.9	677.7	2696.9	1316.9	1987.1	588.6
Mean	2215	631.3	2883.4	1038	3166.1	1239.9	2532.8	650.6

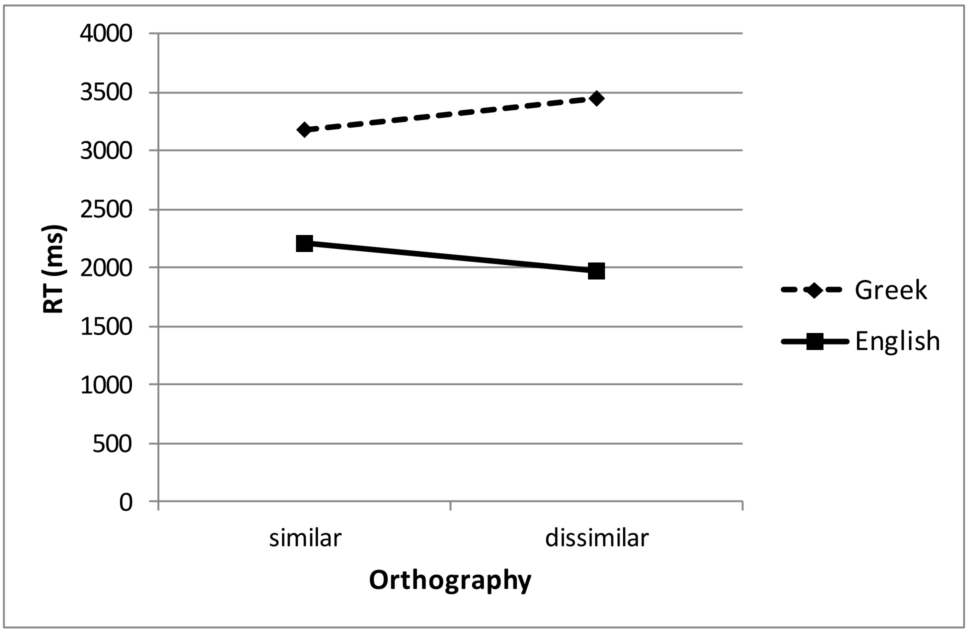


Figure 5.1: Reaction time as a function of Language Group and Orthography for the *Word Rhyme Judgment* task.

Similarly, the interaction between Rhyme and Orthographic Similarity was significant, $F_{1(1,\ 38)}$= 26.7, *MSE* = 16946609.47, p<.001; $F_{2(1,\ 28)}$= 54.19, *MSE* = 212538.85, p<.001. This revealed that Rhyming words that were orthographically similar had shorter RTs compared with orthographically dissimilar words. On the other hand, for non-rhyming words, the reverse was true – i.e. orthographically similar words had longer RTs compared with orthographically dissimilar words (see Table 5.1 and Figure 5.2).

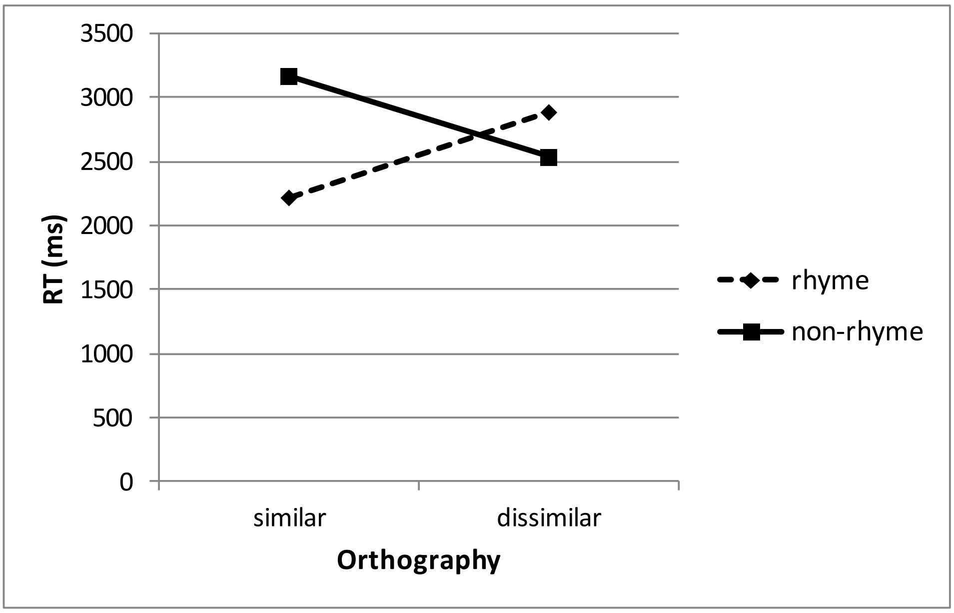


Figure 5.2: Reaction time as a function of Rhyme and Orthography for the *Word Rhyme Judgment* task.

5.3.1.2 RT results for Non-Word Rhyme Judgment task

The RT data for the *Non-Word Rhyme Judgment task* were analysed with a two-way ANOVA (Group x Rhyme). The main effect of Group was significant, $F_{1(1,\,38)}$= 22.44, *MSE* = 49470240.08, p< .001; $F_{2(1,\,48)}$=365.35, *MSE* = 59894520.11, p< .001, showing that RTs for the English participants were shorter than those of the Greek participants (see Table 5.2).

Table 5.2: Mean RTs and SD for Greek and English speakers for the Non-Word Rhyme Judgment Task.

	RhymNonWord		NonRhymNonWord	
	n=25		n=25	
Group	Mean (ms)	SD	Mean (ms)	SD
Greek (n=20)	4095.4	1690	3652.3	1132
English (n=20)	2223.3	740.3	2378.9	742.8
Mean	3159.4	1215.1	3015.6	937.4

The main effect of Rhyme was not significant, $F_{1(1,\ 38)}$= .996, *MSE* = 413126.421, p= .325; $F_{2(1,\ 48)}$= 1.96, *MSE* = 357406.687, p= .168. The interaction between Group and Rhyme revealed significant results, $F_{1(1,\ 38)}$= 4.31, *MSE* = 1792087.102, p<.05; $F_{2(1,\ 48)}$= 12.96, *MSE* = 2363398.902, p<.01. This indicated that the Greek group had longer RTs for the rhyme compared to non-rhyme conditions. However, for the native English speakers, there was little effect of rhyme (see Table 5.2 and Figure 5.3).

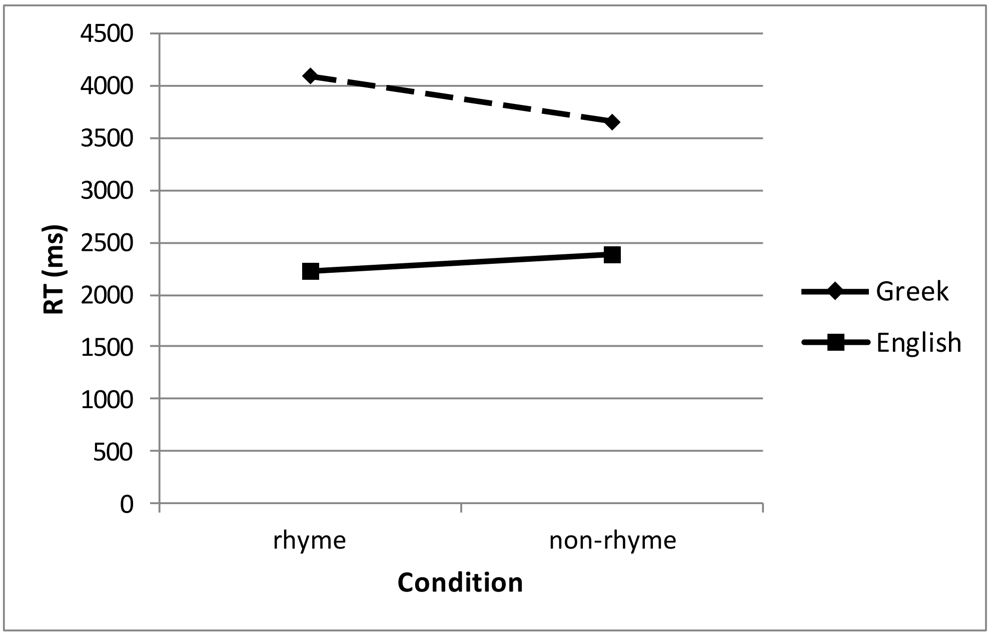


Figure 5.3: Reaction Time as a function of Language Group and Rhyme for the *Non-Word Rhyme Judgment task.*

5.3.1.3 RT results for Homophone Judgment task

The RT data for the *Homophone Judgment task* were analysed with a three-way ANOVA (Group x Homophony x Regularity). The main effect of Group was significant, $F_{1(1,\ 38)}$= 708.56, *MSE* = 72338418.02, p<.001; $F_{2(1,48)}$= 260.2, *MSE* = 24844578.15, p<.001, indicating that RTs for the English group were shorter than those for the Greek group across all conditions (see Table 5.3).

Table 5.3: Mean RTs and SD for Greek and English speakers for the Homophone Judgment Task.

	RegHom n=25		IrregHom n=25		RegNonHom n=25		IrregNonHom n=25	
Group	Mean (ms)	SD	Mean (ms)	SD	Mean (ms)	SD	Mean (ms)	SD
Greek (n=20)	3196.3	1014	3263.7	989.1	2707.3	748.9	2920.4	665.8
English (n=20)	1614.6	408.9	1716.4	290.5	1608	341	1769.6	414.3
Mean	2405.4	711.4	2490	639.8	2157.7	544.9	2345	540

The main effect of Homophony was significant, $F_{1(1,\ 38)}$= 5.06, *MSE* = 1542814.867, p< .05; $F_{2(1,\ 48)}$= 8.69, *MSE* = 2900951.947, p<.01, revealing that RTs for non-homophone words were shorter than those for homophone words (2251.4 ms vs. 2447.7 ms). The main effect of Regularity was also significant, $F_{1(1,\ 38)}$= 8.48, *MSE* = 739575.351, p<.01, revealing that RTs for regular words faster than those for irregular words (2281.6ms vs. 2417.5 ms). The item analysis however did not reveal significant results ($F_{2(1,\ 48)}$= 2.75, *MSE* = 1182493.087, p= .104). The interaction between Group and Homophony revealed significant results, $F_{1(1,\ 38)}$= 6.34, *MSE* = 1931054.738, p<.05; $F_{2(1,\ 48)}$= 8.48, *MSE* = 2832341.707, p<.01. This indicated the Greek group had longer RTs for the homophone compared with the non-homophone condition. However, the English speakers had similar RTs for both homophone and non-homophone conditions (see Table 5.3 and Figure 5.4). No other interactions were significant.

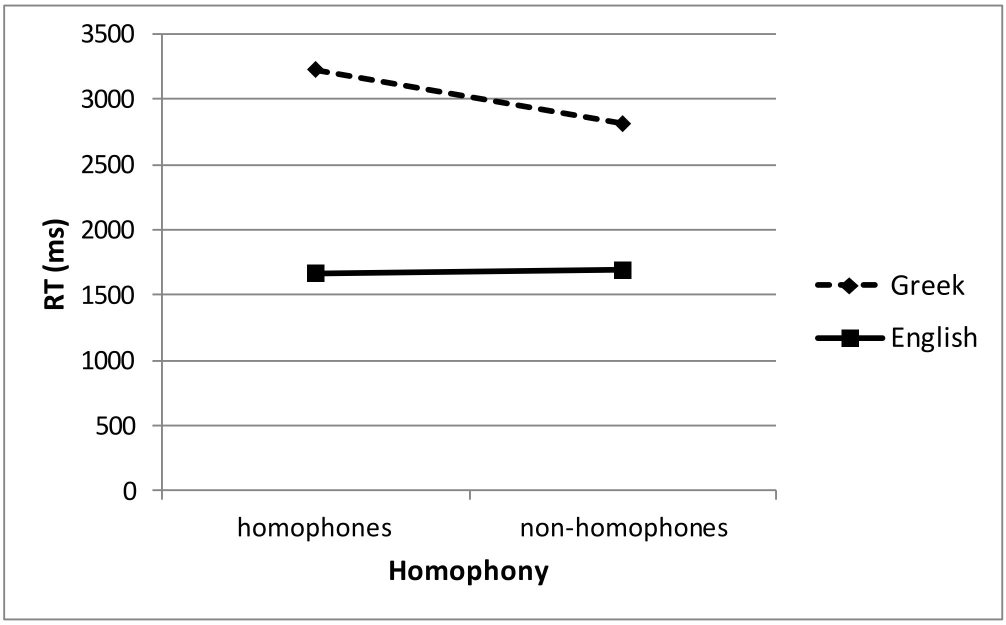


Figure 5.4: Reaction time as a function of Language Group and Homophony for the *Homophone Judgment task*.

5.3.2 Errors

Errors were averaged across participants for the two language groups (Greek and English) for all tasks and conditions. Because errors are not normally distributed, the data were logarithmically transformed after adding a constant of 0.001. This is because sometimes a participant may not have made an error (i.e. a 2% error becomes 0.021, while 0% error becomes 0.001). The error data were then analyzed using mixed—ANOVAs. Table 5.4, 5.5, and 5.6 show the mean error scores and SD for each task and experimental condition for Greek and English groups.

5.3.2.1 Error results for Word Rhyme Judgment task

For the Word Rhyme Judgment task, the error scores were submitted to a three-way analysis of variance (ANOVA Group x Rhyme x Orthographic Similarity). The main effect of Group was significant, $F_{1(1,38)}$= 47.997, *MSE* =

.693, p<.001, showing that the Greek native speakers had a higher proportion of error across all conditions (see Table 5.4).

Table 5.4: Mean Error scores (proportion of error[38]) and SD for Greek and English groups for Word Rhyme Judgment Task.

Group	OrthSimRhym Word n=15		OrthDissRhym Word n=15		OrthSimNonRhym Word n=15		OrthDissNonRhym Word n=15	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek (n=20)	0.3	0.189	0.05	0.07	0.42	0.11	0.05	0.07
English (n=20)	0.04	0.08	0.02	0.04	0.2	0.15	0.03	0.07
Mean	0.17	0.13	0.03	0.06	0.31	0.13	0.04	0.07

The main effect of Rhyme was significant, $F_{1(1, 38)}$= 12.947, *MSE* = .220, p< .01, with non-rhyme words more prone to errors (0.175 vs. 0.1). The main effect of orthographic similarity was also significant, $F_{1(1, 38)}$= 190.085, *MSE* = 1.676, p<.001, with orthographically similar words being more prone to error (0.24 vs. 0.035). There was a significant interaction between orthographic similarity and rhyme ($F_{1(1, 38)}$= 21.308, *MSE* = .194, p<.001), showing that for rhyming words, there was less of a difference between orthographically similar and dissimilar words. However, for non-rhyming words, orthographically similar words were far more prone to error (see Table 5.4 and Figure 5.5).

[38] Measure calculated as a fraction of 1 (i.e. a 0.5 equates to 50% error).

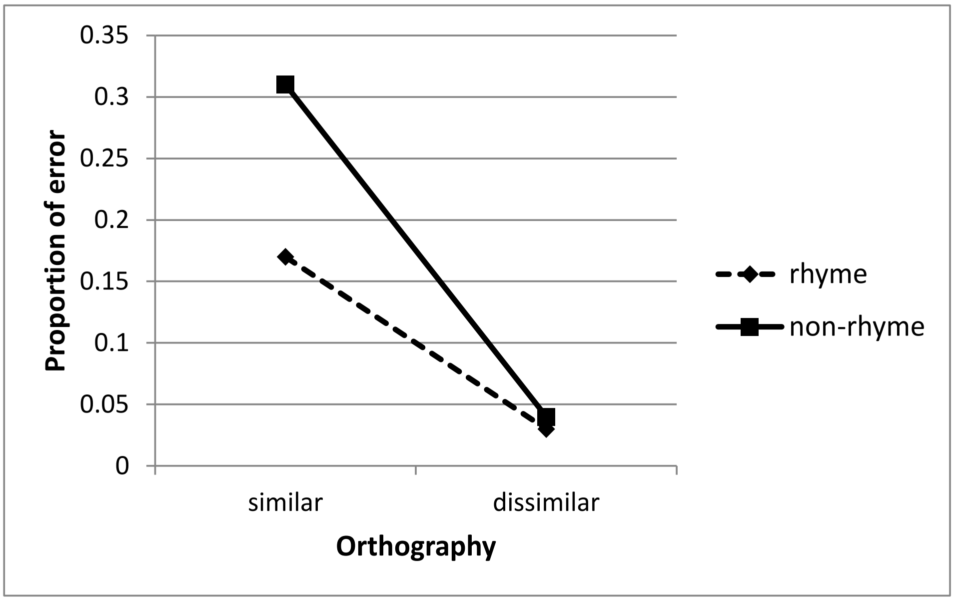


Figure 5.5: Error as a function of Rhyme and Orthography for the *Word Rhyme Judgment* task.

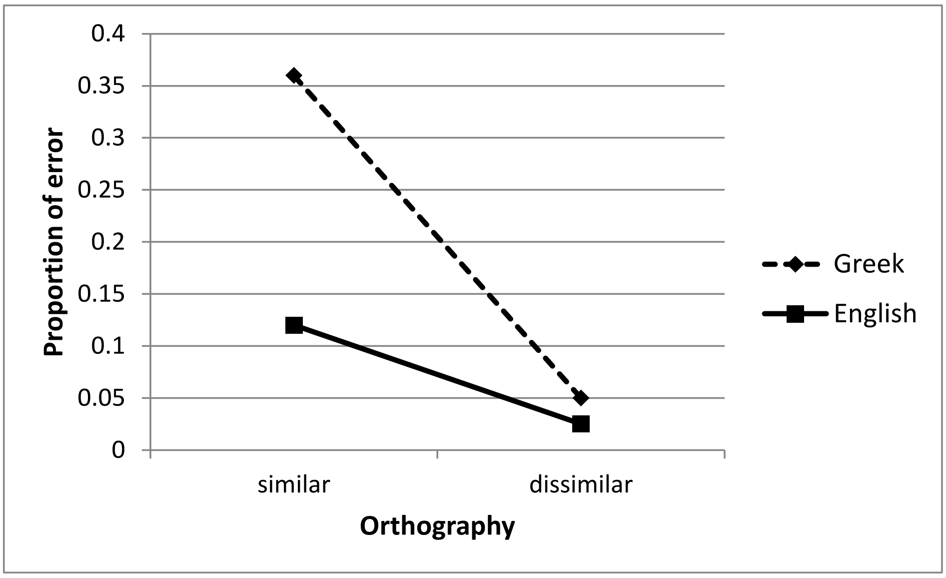


Figure 5.6: Error as a function of Language Group and Orthography for the *Word Rhyme Judgment* task.

The interaction between Group and orthographic similarity was also significant ($F_{1(1,\ 38)}$= 50.920, *MSE* = .449, p< .001). This shows that the Greek speakers were significantly more affected by orthographic similarity compared with

English speakers. Although for both groups, they had higher error rates for orthographically similar items, the effect was more pronounced for the Greek speakers (see Table 5.4 and Figure 5.6). No other interactions were significant.

5.3.2.2 Error results for Non-word Rhyme Judgment task

For the Non-word Rhyme Judgment task, error scores were submitted to a two-way analysis of variance (ANOVA Group x Rhyme). The main effect of Group was significant, $F_{1(1,\ 38)}$= 13.495, *MSE* = .141, p<.01, showing that native language appears to affect reading performance for non-words (see Table 5.5).

Table 5.5: Mean Error scores (proportion of error) and SD for Greek and English speakers for Non-Word Rhyme Judgment Task.

Group	RhymNonWord		NonRhymNonWord	
	n=25		n=25	
	Mean	SD	Mean	SD
Greek (n=20)	0.35	0.16	0.22	0.09
English (n=20)	0.23	0.07	0.17	0.07
Mean	0.29	0.12	0.19	0.08

Similarly, the main effect of Rhyme was also significant, $F_{1(1,\ 38)}$= 15.301, *MSE* = .192, p< .001, with a higher proportion of error for both groups for the rhyming condition (0.29 vs. 0.19). The interaction between Group and Rhyme was also significant ($F_{1(1,38)}$= 1.631, *MSE* = .02, p<.05), with the Greek speakers being significantly more negatively affected by rhyme compared with the English speakers (see Table 5.5 and Figure 5.7).

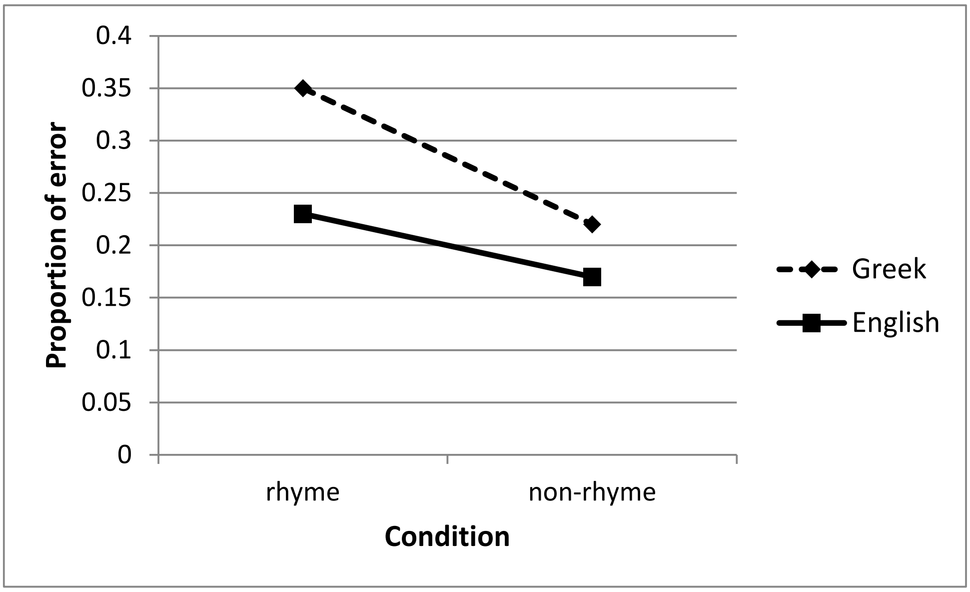


Figure 5.7: Error as a function of Language Group and Rhyme for the *Non-Word Rhyme Judgment* task.

5.3.2.3 Error results for Homophone Judgment task

The error scores for the *Homophone Judgment task* were submitted to a three-way analysis of variance (ANOVA Group x Homophony x Regularity). The main effect of Group was significant, $F_{1(1,38)}$= 29.978, *MSE* = .396, p<.001, with the Greek native speakers having higher error scores across all task conditions. The main effect of Regularity was significant, $F_{1(1,\ 38)}$= 36.705, *MSE* = .110, p<.001 with irregular items having more errors (0.135 vs. 0.085). The main effect of Homophony was significant, $F_{1(1,\ 38)}$= 37.381, *MSE* = .524, p<.001, where the Homophone conditions had higher error rates compared with Non-Homophone conditions (0.17 vs. 0.05). The interaction between Homophony and Regularity was significant ($F_{1(1,\ 38)}$= 11.072, *MSE* = .045, p<.05), such that for non-homophones, there was little difference between regular and irregular items whereas for homophones, there were more errors with irregular items (see Table 5.6 and Figure 5.8).

Table 5.6: Mean Error scores (proportion of error) and SD for Greek and English speakers for Homophone Judgment Task.

Group	RegHom		IrregHom		RegNonHom		IrregNonHom	
	n=25		n=25		n=25		n=25	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greek (n=20)	0.22	0.14	0.31	0.12	0.05	0.06	0.07	0.05
English (n=20)	0.04	0.04	0.11	0.07	0.04	0.07	0.06	0.09
Mean	0.13	0.09	0.21	0.09	0.04	0.06	0.06	0.07

There was also a significant interaction between Group and Homophony ($F_{1(1, 38)}$= 22.33, *MSE* = .313, p<.001) such that for English speakers, they were not affected much by homophony whereas Greek speakers had more errors for homophones (see Table 5.6 and Figure 5.9). No other interactions were significant.

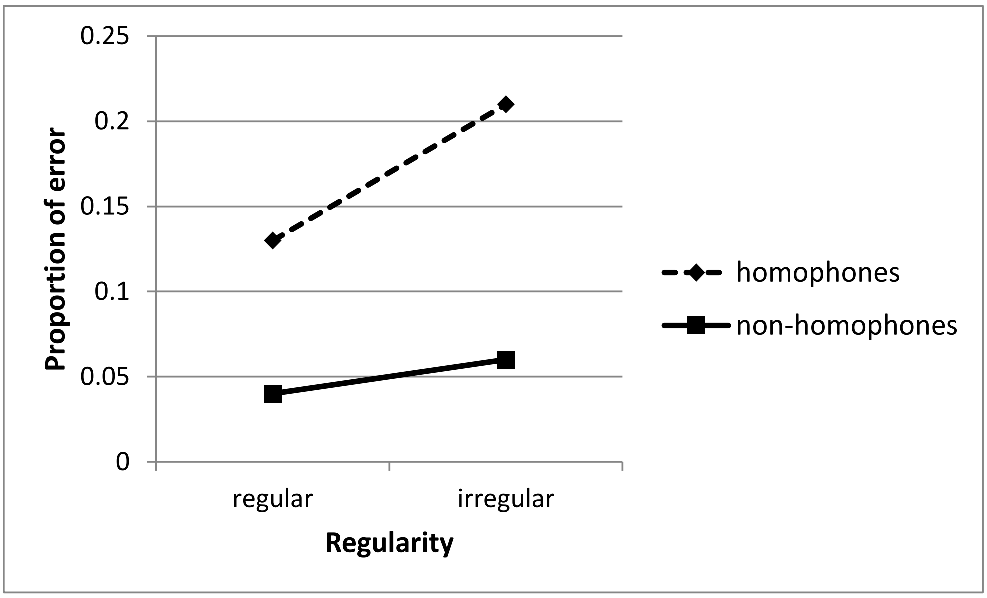


Figure 5.8: Error as a function of Homophony and Regularity for the *Homophone Judgment* task.

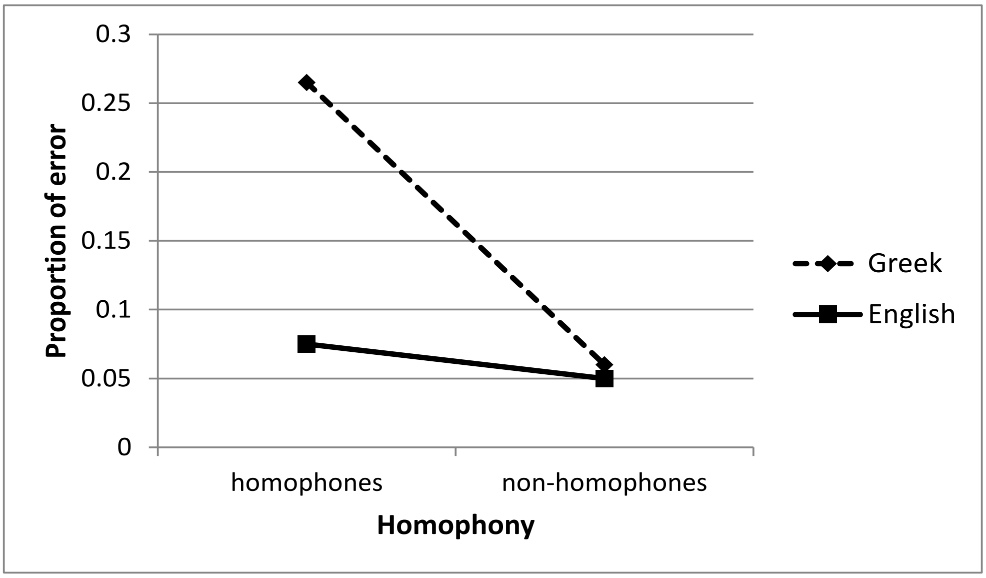


Figure 5.9: Error as a function of Language Group and Homophony for the *Homophone Judgment* task.

5.4 Discussion

This experiment aimed at examining reading processes by adult Greek learners of L2 English in comparison with adult English native speakers in order to investigate how phonological processing could be affecting reading. Specifically, this chapter explored the relationship between orthographic and phonological processes when reading word pairs with different orthographic consistency, and explored the use of orthographic cues by Greek learners of L2 English and native speakers of English. L1 (Greek) for the Greek groups is an orthographically transparent orthography while the target L2 (English) is a rather opaque orthography. This difference in transparency levels between the two languages were expected to result in impaired performance by the L2 group, given the orthographic differences between their L1 vs. L2. Orthographic cues of the L1 (Greek) were also hypothesized to interfere with those orthographic cues offered in the target L2 (English), leading to orthographic transfer. Reading processes were observed with a battery of tests that used word and non-word stimuli, arranged in rhyme judgment and homophone judgment tasks as both tasks require phonological processing. RT (latency) and accuracy data lead to a number of observations.

Results demonstrate that phonological processing influences reading performance and this is mainly due to the consistency (or inconsistency) of the orthographic to phonological mappings between the L1 (Greek) and the target L2 (English). Specifically, the lexical decision tasks that involved orthographically similar rhyming word pairs were faster and less error prone than in the case of orthographically dissimilar rhyming word pairs. This provides evidence that orthographic similarity can affect L2 learners' reading performance and it also supports earlier findings by Seidenberg & Tanenhaus (1979) that two similarly spelt words (e.g. *hear – near*) were recognized faster by participants compared to dissimilarly spelt words (e.g. *hall - bawl*) (see also, Tanenhaus, Flanigan, & Seidenberg, 1980; Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981). This result, however, was not confirmed in the case of non-rhyming word pairs with orthographic similarity (e.g. *weight – height*) as predicted and this could be attributed to the fact that such orthographic similarity can be a misleading 'cue' as in this case the two words do not rhyme. Therefore, orthographic similarity seems to aid reading performance for rhyming words, but is causing interference for non-rhyming words. This outcome also suggests L1 transfer: Greek being a transparent language does not allow for such phonetic variation as realized in English for cases such as e.g. *weight* vs. *height*, where the same letter clusters *'-eight*' can be pronounced in two different ways. Therefore, this also supports the argument of L1 orthographic transfer to L2 reading. In addition, the hypothesis that shared orthography effects can inhibit correct reading performance due to already existing mappings of grapheme-to-phoneme correspondances (e.g. Pytlyk, 2012; Oller & Ziahosseiny, 1970) is also supported.

This explanation is also congruent with or analogous to the findings in the previous experiments described in Chapter 2, 3, and 4, particularly showing that vowels do offer cues which could be used both at an acoustic / phonetic level as well as the orthographic level. The results of this experiment also complement the findings of Experiment 2 (Chapter 3) where the orthographic word stimuli were replaced with picture stimuli which revealed that when orthographic cues are removed (or not available) this creates impairements in the perceptual identification of the target vowels. Experiment 2 and 5 therefore lead to the assumption that orthography offers cues that affect correct identification of English words for the L2 learner and also further support the existing relationship between orthography and phonology (for L2 sound perception) (e.g. Erdener & Burnham, 2005; Hallé *et al*., 2000).

This study also suggests that factors such as orthography and rhyme seem to jointly affect reading processing and therefore appear to be associated with phonological processing. For example, orthographically similar word pairs differed in accuracy results depending on rhyme or non-rhyme conditions. Rhyme vs. non-rhyme factors seem to also affect RT performance. This was particularly obvious for non-words where L2 learners had longer RTs for the rhyme than non-rhyme condition which was also confirmed with accuracy scores. This outcome suggests a perceptual conflict between orthographic similarity and rhyme for the Greek learners of L2 English. This can be explained in view of the transparent Greek orthography and an 'expectation' that orthographic similarity governs rhyming effects which is not necessarily true in English.

Another observation is that of homophony where a correct response would require access to the full phonological code (see e.g. Jared & Kroll, 2001; Jared & Szucs, 2002) as, for example, in the case of *blew* – *blue* (homophones) versus *talks* – *tax* (non-homophones). Homophone words had longer RTs and were more error prone than non-homophone words, and homophony was also observed to interact with regularity but only in the case of irregular homophone words. The case of irregularly spelt homophones that caused confusions to the L2 learners leads to the assumption that their L1 phoneme-to-grapheme mappings were 'violated' when reading irregular spellings of L2 English word pairs despite their identical sound (i.e. homophone pairs).

This also relates to the 'shared orthography' effect suggested to cause perceptual confusions (e.g. Oller & Ziahosseiny, 1970) since in order to overcome any such confusions, the L2 learner will need to disassociate the grapheme-phoneme accosiations that already exist in their L1 (e.g. Jared & Szucs, 2002; Treiman & Cassar, 1997). Therefore, the L2 learner needs to re-arrange these L1-L2 phoneme-grapheme mappings in order for correct L2 reading. In this study however, despite the amount of years in L2 instruction, Greek participants seem to still encounter more difficulty for reading specific word pairs (e.g. irregular homophones) compared with others (e.g. regular homophones). This observation seems to support the proposal that, once learnt, speech sound and spelling are hard to be separated (see e.g. Treiman & Cassar, 1997; Jared & Kroll, 2001; Jared & Szucs, 2002) and that L1 phonological representations are automatically activated in L2 reading (e.g. Jared & Szucs, 2002). This 'unwanted interference' (Landerl *et al.*, 1996, p.12) has been attributed to the consistency of orthographic and phonological mappings between two languages (e.g. Schwartz, Kroll & Diaz, 2007) that influence reading processing for the L2 learner.

The fact that English participants performed better across all tasks compared to Greek participants could lead to two obvious explanations: 1. The native English speakers would obviously have far more reading experience with the target language (i.e. English) compared with the Greek L2 learners of English; 2. The degree of orthographic transparency of the target language (i.e. English), depending on whether it is an L1 or L2 for the reader. For Greek native speakers, reading L2 English could be described as a cognitively more demanding task (see also Pytlyk, 2012) because their L1 (Greek) is a transparent orthography while the L2 (English) is a less transparent (opaque) orthography. This result seems to be in line with the Landerl *et al.* (1997) findings discussed earlier that consistency in grapheme-to-phoneme relations essentially aid readers, even readers with phonological deficits, to learn print-to-sound mappings but opposite results are observed when these mappings are less consistent or less transparent (i.e. opaque).

This result, however, reveals two significant limitations with the design of this study: 1. the first explanation leads to an assumption that may confound the amount of practice or skill level when Greek learners are compared with

English native speakers. To fully answer this question it would be necessary to additionally compare these results with learners of L2 English who have an opaque L1 language (e.g. French); 2. the second explanation postulates that in order to be able to compare one-to-one versus one-to-many grapheme-phoneme correspondence, it would be necessary to also test an additional control group with an opaque L1, other than English, for example French. This way it would be possible to compare both L1 transparent (e.g. Greek) to an opaque L2 orthography (e.g. English) and also L1 opaque (e.g. French) to an opaque L2 orthography (e.g. English). Such design would allow for a more complete discussion on grapheme-phoneme correspondence in relation to language transparency as well as how different transparency levels of the L1 (i.e. transparent vs. less transparent /opaque) could affect reading performance in the opaque L2. This limitation in the experimental design of this study allows only for partial assumptions in that L1 Greek appears to affect reading performance in L2 English and this can be attributed to grapheme-phoneme correspondences that interfere between a transparent L1 versus an opaque L2 orthography. Further tests with an experimental design that would aim to overcome the limitations identified in this case could provide more conclusive assumptions.

Assessing the experimental data of this study in relation to the reading models discussed in Chapter 1, namely the '*Dual-Route Cascade model*' (DRC) (Coltheart *et al.*, 2001; Coltheart & Rastle, 1994; Rastle & Coltheart, 1998), and the '*Parallel Distributed Connectionist*' (PDP) reading model (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996; Plaut & Kello, 1999), a few observations can be made.

The PDP model predicts that lexical decision is more difficult in the case of non-words that are similar to real words (as in the case of the current experiment) compared to non-words that are dissimilar to real words. This effect has been explained (see Plaut, 1997) based on the fact that more difficult lexical decision tasks require semantic representations. A recent study by Bormann & Weiller (2012) has investigated the case of a 'word-meaning deafness' patient, a quite rare cognitive impairment which, however, gives light to issues such as the need of a mental 'lexicon' (or a semantics unit in the context of the PDP model) for the case of lexical decision tasks. Bormann & Weiller (2012) found that their participant could still read non-words similar to real words with comparable accuracy. A main outcome of the Bormann & Weiller (2012) study is that lexical decision could not be based on semantic information (since in this case it was not available) but they assume the existence of lexical representations (in other words, a mental lexicon).

In view of the data in the current study, English native speakers had no particular difficulty with any word /non-word rhyme judgment or homophony task as confirmed by accuracy and latency scores. On the other hand, for Greek L2 speakers of English accuracy and latency scores were particularly impaired across conditions but not across tasks. This is a particular implication for the PDP model as L2 learner outcomes would have to affect task (e.g. non-word rhyming judgment task) rather than specific conditions within the task. For example, results showing that rhyme conditions had larger error scores for the L2 learners compared with the non-rhyme condition imply that phonological and generally 'lexical' processing does not depend on semantic processing of the stimuli. The PDP model, however, cannot provide a complete account for such outcome as in this case the semantic unit could even be proposed as redundant.

Abandoning a semantic unit could be proposed as a rather radical proposal however, recent research seems to favor the existence of independent lexical representations (e.g. Coltheart, 2004, 2006; Miozzo, 2003; Miozzo & Gordon, 2005). The current data could support Bormann & Weiller (2012)'s findings mentioned earlier in a way that a semantic component would not be necessary for word/non-word rhyme judgment tasks but lexical representations would be necessary. These lexical representations in the case of L2 learners, however, could be interfering between the L1 and the L2, in the form of grapheme-phoneme correspondances and 'violation' of already established L1 representations conflicting with the L2. Some theorists have connected such effects with readers of transparent orthographies, even suggesting that the consistency of grapheme-phoneme could lead to reading processes not occurring via a semantic route (e.g. Allport, 1979; Katz & Feldman, 1983; Besner & Smith, 1992; Katz & Frost, 1992); and, the more extreme views even suggesting that reading via a semantic route never develops in transparent orthographies (e.g. Turvey *et al*., 1984; Bridgeman, 1987) as readers can read through simply following grapheme-to-phomeme conversion rules. The shared orthography effect could also be a plausible explanation in view of the current data rather than an explanation based on PDP predictions.

In contrast, the data appear easier to reconcile with the DRC framework, a dual reading route that assumes localist representations for lexical and sublexical knowledge. The DRC model postulates the existence of two parallel routes being activated when reading, a lexical route that accounts for reading real words and requires the existence of a mental lexicon and a phonological route which can account for the reading of non-words or novel words. It also proposes a 'lexical lookup procedure' (Rastle & Coltheart, 2000, p. 343) for correct articulation, especially for regularity or rhyme conditions, and a possible 'conflict' between the lexical and sub-lexical routes can lead to longer reaction times. Data of this study can confirm these predictions, by the fact that longer latencies (for example, in the case of orthographically dissimilar words) did not necessarily result in error. This confirms the time required for a 'lexical lookup' (Rastle & Coltheart, 2000, p. 343) which gives longer RT scores.

Also, non-words are usually expected to have longer RTs. Slightly longer RTs were observed in the experimental results for non-word as opposed to word rhyme judgment tasks. According to the DRC model results are interpreted based on the assumptions of the grapheme-to-phoneme (GPC) route for reading (Rastle & Coltheart, 1998) as the lexical route cannot aid correct pronunciation for non-words (e.g. Wydell *et al*., 2003), although a semantic 'look up' procedure would still be required. Although the DRC model would expect orthographically similar rhyming pairs to produce faster RTs compared to orthographically dissimilar rhyming pairs, results do confirm that orthographic similarity (or consistency) affects RTs accordingly, however, rhyming effects also influence RTs as already discussed. The DRC can provide a more plausible explanation for L2 learners' impaired performance affected by rhyming or regularity conditions for both words and non-words. Generally, it appears that the DRC can account for more variance in word and non-word naming, unlike the PDP.

The present experiment is an initial exploration investigating phonological processing in tasks involving reading, with an emphasis on the existing relationship between orthography and phonology, specifically Greek as L1 and English as the target L2. It is evident that results presented here are suggestive rather than conclusive. However, these initial findings depict interference between the L1 and L2 orthographic systems (based on transparency levels for each of these languages) for the L2 learners considered here. It is evident that further tests would be necessary in order to make more concrete explanations in terms of how specific orthographic factors of a L1-L2 relationship affect phonological representations and therefore reading in an L2. Interpreting this experiment in relation to experiments presented and discussed in previous chapters (chapter 2, 3 and 4), a main assumption is that speech perception outcomes for the L2 learner may not be entirely based on a phonological level based on the L1-L2 phonological relationship, but the respective L1-L2 orthographic relationship may also contribute to some extent to phonological representations and essentially perception of L2 speech. Specifically, there may be orthographic cues used by learners of L2 English as a 'mechanism' for L2 reading. English being an opaque orthography, however, appears to cause distortions for the L2 readers when using such 'mechanism' due to differences in grapheme-phoneme correspondences between Greek (L1) and English (L2). These preliminary results aim to point out this issue and are suggestive for further investigations.

Future experimental extensions could include both words and non-words in the same task (as a mixed design) in order to be able to identify any significant RT changes based on lexicality effects. An additional interesting extension could also be the testing of child participants as in the case of previous experiments discussed in earlier chapters in order to additionally test maturational effects. The age group of child participants should however take into account the L2 English curriculum in order for the word stimuli used in experimental tasks to be already taught at the school level of the target age group. It is also important for child participants to be familiar with concepts such as ‘homophony’ and ‘rhyme’. In future, it would be interesting to include more language groups (see also an interesting account by Borgwaldt, Hellwig, & De Groot (2005) on letter-to-phoneme mappings in 7 languages). Such extension would allow a more powerful design in order to compare transparent versus opaque L1 orthographies (e.g. Greek and French) with an opaque L2 orthography (e.g. English) that would account for the current limitations.

Chapter 6

General discussion and conclusion.

Given the fundamental role of language for cognition and communication, it is not surprising that the foundation for this ability is shaped in the early stages of development. Not only speech sounds are preferentially processed at birth over other non-speech sounds (e.g. Telkemeyer, Rossi, Koch, Nierhaus, Steinbrink, Poeppel, Obrig, & Wartenburger, 2009; Vouloumanos & Werker, 2007), but there is growing evidence of sensitivity to speech sounds even pre-natally (Kisilevsky, Hains, Brown, Lee, Cowperthwaite, Stutzman *et al*., 2009). These predispositions are obviously the basis for relatively easy language acquisition during the first year of life (Gervain & Mehler, 2010) and beyond. These language abilities have received much attention over the past few decades with particular emphasis on infants' initial remarkable language capabilities describing them as 'citizens of the world' (Werker & Tees, 1984; also, Eimas *et al*., 1971). A decline, however, in these language acquisition abilities even from the early developmental stages (e.g. Kuhl *et al*., 1992; Polka & Werker, 1994), and the degree and ease of learning of second and/or additional language(s) are central issues in ongoing research (e.g. Escudero *et al*., 2008; Broersma, 2005; Ylinen *et al.*, 2008; Lengeris & Hazan, 2010, to name but a few). This thesis sought to provide an empirical contribution to the fields of speech perception and L2 learning in a specific population: Greek learners of English. It examined English vowel perception and cue weighting, specifically by Greek learners of L2 English, highlighting age effects and maturational constraints. More specifically, an issue addressed here is the identification and discrimination of L2 speech sounds and how the L1 could be affecting cue weighting. The possibility of orthographic cues also contributing and affecting speech perception was also tested in view of a more specific examination into cognitive processes of reading. This thesis also emphasises the beneficial contribution of perceptual training techniques (e.g. the HVPT), and discussed the didactic potential of such techniques in light of the prominent speech learning models of L2 acquisition.

6.1 L2 Vowel Perception, Cue weighting and Age effects: "It is all Greek to me"

The learning of a L2 speech sound by a L2 learner involves the formation of new phonetic categories which also includes weighting of phonetic cues that are often weighted differently between the target L2 and the learner's L1 (e.g. Strange *et al*., 1998; Flege, 1988; Flege *et al*., 1997; McAllister *et al*., 2002; Ylinen *et al*., 2010; Aoyama *et al*., 2004; Bion *et al*., 2006; Rauber *et al*., 2005; Escudero *et al*., 2008; Broersma, 2005; Cutler & Broersma, 2005). Therefore, those cues that are critical for a native speaker of a language could be weighted as secondary by a L2 learner of that language (e.g. Holt & Lotto, 2006). Elements of the L1 (e.g. such as cues and other language components that may be critical for speech sound identification) could transfer (Bohn, 1995; Strange, 1998) and thus restrain L2 speech perception. In this sense, L2 learners may rely on secondary cues in perceptual categorization tasks rather than the relevant and critical cues used by native speakers of the target language (e.g. Iverson *et al*., 2003).

Experiment 1 tested the perceptual identification and discrimination of the English tense-lax (/iː/ - /ɪ/) distinction by Greek learners of L2 English. The use of natural and modified duration stimuli allowed measuring the reliance on spectral and/or duration cues. It was clear that for the case of Greek speakers of L2 English, L2 cues were weighted differently by L2 learners in comparison with the L1 groups. Perceptual interference effects between L1 and L2 phoneme categories were observed whereby L2 learners' preference to specific phonetic cues (in this case duration over spectral cues) that are not used as primary cues by the L1 speakers (as revealed in the case of modified duration stimuli). The outcome for the L2 learners was that accurate identification of the /iː/ - /ɪ/ contrast under consideration was impaired, and this was particularly pronounced when modified duration stimuli were used (which forced reliance on spectral cues). This observation is in line with Best's PAM/PAM-L2 (Best, 1995a, Best & Tyler, 2007) proposal that L2 categories assimilate into L1 categories. In this case, the two distinct and contrastive English vowel categories of /iː/ and /ɪ/, seem to be assimilated into a single /i/ category which exists in the Greek vowel inventory (see Chapter 1, also Lengeris, 2009) and duration appears to be primarily used by L2 learners as a cue for L2 vowel categorization.

Experiment 1 suggested that duration appears to be an easier cue to rely on (e.g. Ylinen, Shestakova, Alku, & Huotilainen, 2005; Escudero, 2005; Cebrian, 2006). A main assumption therefore is that Greek learners of L2 English weight cues differently in the L2 and vowel identification and discrimination is based on different cues (i.e. duration cues) to those weighted by English native speakers (i.e. spectral cues) (also, Escudero, 2005; Cebrian, 2006; Legeris, 2009, among others). In this way, this sensitivity to duration cues by the Greek groups supports the desensitization hypothesis (Bohn, 1995), which proposes that L2 learners can have access to duration cues even if not used contrastively in their L1. The 'feature hypothesis' (McAllister *et al*., 2002) which proposes that L2 learners do not have access to L2 cues (or features) that are not used contrastively in their L1, is not supported. The fact that duration is not a phonemically relevant cue in Greek but Greek vowels are realized by different place of articulation (as discussed in Chapter 1), could lead to the prediction that Greek participants would be perceptually more sensitive to spectral rather than duration cues. However, the modified duration condition showed an over-reliance on duration cues by the Greek participants. This over-reliance on duration cues for the L2 English /iː/ - /ɪ/ identification and discrimination suggests that it is still possible for such a secondary cue to be employed and that cue weighting in the L2 cannot always be predicted.

The assumption that duration is merely an ‘easier’ cue for L2 learners has been challenged by a recent study by Lengeris (2009) which suggests that duration is not entirely or exclusively weighted by L2 learners in relation to spectral cues. Lengeris (2009)’s study looked at between-context comparisons for the different English vowels (based on nine contrastive vowel pairs) and found equivalence in duration between L1 and L2 categories even though there may be no phonemic vowel length contrast in the L1. This suggests that there is assimilation of both duration and spectral cues of the L2 vowels into L1 categories and that duration may not be used by L2 learners exclusively within the context of the desensitization hypothesis (Bohn, 1995) as previously proposed. Therefore, although the outcomes of Experiment 1 which focused on the English /iː/ - /ɪ/ contrast do suggest access to duration cues by the L2 learners and duration being weighted as a critical rather than a secondary cue, it is important to note that more L1 vs. L2 contrasts should be tested in order to achieve a more holistic approach to cue-weighting and speech sound category assimilation patterns. Nevertheless, Experiment 1 is in support of the prominent L2 speech perception models (see Table 6.1), namely Best (1995a, Best & Tyler, 2007)’s PAM/PAM-L2 and Flege (1995a)’s SLM, which highlight L1 transfer, as well as the PI account (Iverson *et al*. 2003; Kuhl *et al*., 2006, 2008) that suggests L2 learner’s sensitivity to irrelevant (or non-critical) acoustic variation as a result of L1 experience. Native language magnet effects are also supported as the L1 seems to affect L2 speech perception in way that native speech categories already perceptually formed seem to be powerful attractors of L2 speech sounds, as the NLM/NLM-e (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995; Kuhl *et al.,* 2008) proposes. This mechanism could be suggested to ‘aid’ perceptual assimilation patterns for the L2 learners. Overall, it appears that L2 learners do transfer L1 patterns which can aid or impair perception of L2 speech sounds also depending on duration relationships between the L1 and L2.

Considering de Jong & Zawaydeh (2002)’s proposal that consonants surrounding a vowel could generally affect the vowel quality, Experiment 3 aimed to explore and control for the existence and possible contribution of perceptual cues carried by the consonant frame (C_C) surrounding the vowel pairs. The consonants surrounding the vowel contrasts under consideration were cross-spliced (e.g. Kemps *et al*., 2005) which allowed for observations specifically for the minimal pairs under consideration. Experimental results mirrored those of Experiment 1 which supported the assumption that there was no interference between the C_C frame and the contrasting vowel sounds of the minimal pairs used. Therefore, the C_C frame cannot be considered as carrying information that could be perceptually identified and weighted as important cues toward vowel identification and discrimination within this specific experimental context. An explanation therefore on vowel perception and cue weighting is to be sought within the vowel contrast itself.

The testing of different age groups (Experiment 1) allowed for observations on age-related effects and maturational constraints, in relation to speech perception models. First, English native speakers had high accuracy scores regardless of their age: high accuracy scores were achieved by 7-8, 9-10, 11-13 year old English child participants as well as 20-30 year old English adults. This pattern of maturational changes is in line with Kuhl's (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995; and, Kuhl *et al*., 2008) NLM/NLM-e model that the native language acts like a perceptual magnet and emphasizes these native language effects taking place from birth and native speech categories perceptually forming within the infant's first year of life (e.g. Kuhl, 2004). Therefore, native speech categories have already been established not only for the English adult native groups but also for the English child native groups, aged 7 to 13 years of age.

The impaired performance of all three Greek child groups also is a reflection of Kuhl's (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995; and, Kuhl *et al*., 2008) NLM/NLM-e model, but in this case the powerful native attractors are speech sounds of the native Greek language. This not only seems to impair the Greek child participants' performance but it also strengthens the assumption that English speech sound pairs that are normally contrastive, have been perceptually assimilated by the L2 (Greek) learners. In this case, Best's (Best, 1995a; Best *et al*., 2001, Best & Tyler, 2007) PAM/PAM-L2 model, predicting this outcome, is supported.

Age effects, however, were observed between the L2 child and adult groups. It appears that the years of L2 experience that adults had is a contributing factor for better performance (accuracy scores). As proposed by the SLM (Flege, 1995a, 2003) it is easier to identify L2 speech sounds for the L2 learner following years of experience with the L2. Of course, Flege's suggestion comes primarily from populations that had actually immigrated to the country where the target language was spoken (e.g. Flege & Murno, 1994; Fox, Flege, & Murno, 1995; Flege & Liu, 2001), whereas in the case of this thesis, Greek participants were learners of L2 English in their own native country. Experiment 1, 2 and 3, discussed the impaired performance across tasks of L2 child participants compared with L2 adults. L2 children's generally low scores (especially for the perceptual identification tasks), led to the suggestion that there may be implications of both age and proficiency, since L2 child and adult participants differed at both age and proficiency levels. Given that L2 adult participants had received more years of L2 instruction compared with L2 child participants who were at beginner levels of L2 learning. Therefore, results could not be solely attributed to an over-reliance on duration cues, or be explained based on merely age effects. Even though this was a plausible explanation, the experimental design did not allow to directly test these assumptions in order to see which explanation could be more likely. However, these results generally support evidence provided by Flege & Liu (2001, p. 527) who relate the amount of L2 input with the fact that 'adults' performance in an L2 will improve measurably over time'. Nevertheless, a main assumption from experiment 1, 2 and 3 could be that age effects between adult and child L2 learners' task performance could be the outcome of an interaction with the difference in proficiency levels between the two age groups. However, this assumption was challenged in Experiment 4 discussed in section 6.3.

6.2 Hearing or 'reading' the sounds?

The possibility of interference between orthographic cues and perceptual identification of L2 speech sounds was explored in Experiment 2. Following translation and picture match tasks to ensure knowledge of the picture stimuli and respective 'meanings' used throughout Experiment 2, the main outcome confirms reliance on duration cues, as already observed in Experiment 1, and also introduces a new dimension to a possible explanation: speech perception in relation to spelling (or orthography) of words. It was hypothesized that it is possible for words to carry important orthographic cues which could aid perceptual identification of the L2 vowel contrast (/iː/ - /ɪ/). Experiment 2 showed that for the L2 learners the use of pictures instead of words had an effect in their performance. In this case, pictures seemed to 'hinder' accurate identification of speech sounds while words appeared to 'help'. This gives rise to the proposal that L2 learners could be 'reading' the sounds besides 'hearing' the auditory input.

Given that this was just a single study, results cannot be exhaustive but rather suggestive. This is the first study (to the author's knowledge) to test perceptual identification using picture stimuli therefore literature is scarce. Previous studies have mainly concentrated on dyslexic populations and have attempted to link phonological deficits with interferences between spelling and sound (e.g. Liberman & Shankweiler, 1985; Rack, Snowling, & Olson, 1992; Stanovich, 1988; Wagner & Torgesen, 1987; Bradley & Bryant, 1983; Bruck, 1992; Manis, Custodio, & Szeszulski, 1993), but no study has previously tested non-impaired participants in a perceptual identification context with picture stimuli. It is therefore not possible to compare these results with similar previous studies. However, it is fair to say that the observed patterns suggest not only reliance on duration cues (as revealed by the natural versus modified duration tasks which also mirror the results of Experiment 1) but also reliance on orthographic cues since the perceptual identification task using word stimuli yielded higher accuracy scores compared with the same task which involved picture stimuli (see Table 6.1). Future studies using a wider number of stimuli and participants would be necessary in order to confirm and extend the present outcomes.

6.3 Training of vowel categories and Age effects reversed.

A number of previous training studies have focused on training vowel contrasts, for example, training Japanese speakers (e.g. Lambacher *et al*., 2005; Nishi & Kewley-Port, 2007), Korean speakers (e.g. Nishi & Kewley-Port, 2008), German and Spanish speakers (Iverson & Evans, 2007; Iverson & Evans, 2009), and Finnish speakers (Ylinen *et al*., 2010). One previous training study (Lengeris & Hazan, 2010) has considered individual variability in L2 vowel learning, more specifically, Greek learners of L2 English. Lengeris & Hazan (2010) used a large number of pre- and post-training tests, natural and synthetic vowels in both English and Greek, and tested Greek participants' vowel perception, vowel production, and frequency discrimination, in quiet and noisy experimental conditions. Lengeris & Hazan (2010)'s study is a highly relevant and significant contribution to the body of research presented here, particularly due to the L1 populations considered. However, a novel empirical contribution is offered here in the additional training of child participants which is in fact an important extension to previous perceptual training studies in general.

Participants' perceptual identification and discrimination of the English /iː/ - /ɪ/ vowel contrast following high variability perceptual training (HVPT) intervention, was explored in Experiment 4. This study included 10 training sessions for both adult and child participants and included training using natural duration stimuli for half of the participants and modified duration stimuli for the second half of the participants. The aim was to observe age-related training effects and the outcomes of targeted perceptual training as well as to explore the link between the brain's early learning capacity and L2 experience, following perceptual training intervention that has already been a successful intervention in previous training studies (e.g. Ylinen *et al*., 2010).

Also, by training half of every age group with modified duration stimuli it was possible to examine whether those groups could re-learn cue-weighting in a native-like manner (i.e. attending primarily to spectral rather than duration cues).

Pre-training tests confirmed perceptual interference between L1 and L2 vowel categories (Iverson *et al.*, 2003) and suggested L2 category assimilation (Best, 1995a). Pre-training results also mirrored those results obtained in Experiment 1. Accuracy was impaired for the L2 learners, especially in the case of modified duration tasks where perceptual identification and discrimination was necessary to be based solely on spectral cues. English native speakers had high accuracy scores and age was not an affecting factor. L2 learners performed at chance levels, adults slightly better than children and modified duration tasks had low accuracy scores compared with the natural duration tasks. Duration cues again seemed to be used by the L2 learners for vowel distinction of the /iː/ - /ɪ/ contrast.

High variability perceptual training improved participant's post-test accuracy scores and cue-weighting improved considerably. Although the general aim was for the training intervention to reflect natural speech, unlike modified duration stimuli, results show that manipulation of duration cues helped L2 learners to perceive spectral cues that were already naturally present in the /iː/ - /ɪ/ contrast. This kind of targeted training could be considered as a successful approach to L2 learning, given the experimental outcomes: L2 learners who were trained with modified duration stimuli had higher accuracy scores in the post-training tests with relatively more improvement specifically for the modified duration stimuli tasks, especially considering the pre- and post-training difference. Although a larger participant sample would ideally be necessary in order to lead to more definite conclusions, the acquired data presented in Experiment 4 suggest that targeted training (such as modified duration stimuli) could offer 'better' training in terms of specifically 'leading' perception toward more native-like cue weighting and avoiding perceptual interference with non-critical cues. This issue would require further testing in future studies.

This data also support the proposed ‘attention to dimension’ (A2D) models of speech perception (e.g. Iverson & Kuhl, 1995, 1996; Goldstone, 1993, 1994; McClelland, 2001; Nosofsky, 1986) which highlight that it is possible to reallocate L2 learners’ attention to the ‘relevant’ dimension (cue). This proposal is supported not only by the targeted training for those L2 groups trained with modified duration stimuli, but also applies to the HVPT approach in general as shown in the results of Experiment 4. Perceptual training appears to enable the shaping of L2 learners’ perceptual space and the shifting of attention to relevant dimensions (e.g. Francis & Nusbaum, 2002; Kuhl & Iverson, 1995; Jusczyk, 1989). Data suggest that cue-weighting can be re-learnt and targeted stimuli type can also be suggested to aid such perceptual re-arrangement.

An additional important outcome is the degree of improvement for each group following the training intervention which point to maturation constraints. Age differences are observed, however, post-training results show reversed outcomes to those observed in pre-training tests or even the trends of Experiment 1 between adult and child groups. HVPT improved perceptual identification and discrimination for L2 vowel contrasts, but the degree of improvement for child groups, however, was more pronounced. When comparing the difference scores between adult and child groups, it is clear that, children overall improve more following HVPT. In other words, despite the lower starting points and the comparatively less L2 experience for the child groups compared with the adult groups, improvement is significantly larger for the child groups. It could thus be claimed that, following the training intervention, age-effects are reversed and differences in L2 experience (or proficiency level) between the two age groups do not seem to affect learning trends. Experiment 4 challenged earlier assumptions of age interfering with L2 experience or proficiency levels. Given the children’s marked post-training improvement it could be proposed that proficiency level could not be suggested as restricting the degree of improvement following perceptual training, particularly in view of the L2 children’s lower baseline levels. On the one hand, this outcome merely shows adults’ neuroplastic capacity for such adaptation to occur which also stretches the boundaries proposed by critical and sensitive periods for language acquisition (e.g. Chiswick & Miller, 2008; Hyltenstam & Abrahamsson, 2003; Scovel, 1988, 2000, 2006; Long, 1990; Johnson & Newport, 1989). But, on the other hand, the fact that child groups have the most pronounced improvement demonstrates: a. children's high neuroplastic capacity for learning, b. the potential of such perceptual training techniques for educational purposes.

A final observation is this of individual differences when looking at individual data from the course of the 10-day training period. Although there was variability in terms of individual start points, end points as well as the in-between stages of the training period, there is also a general consistency in terms of clear learning trends. Those learning trends were observed for all age groups, but child groups showed a more robust learning curve which also confirms children's susceptibility to such plastic changes and specifically re-arranging of phonetic categories. This outcome supports previous studies that have also emphasized individual differences, even following training intervention (e.g. Lengeris & Hazan, 2010; Hazan *et al*., 2005, 2006; Golestani & Zatorre, 2009; Lee *et al*., 2007; Bradlow *et al*., 1999).

The outcomes of this study are also compatible with the prominent speech learning models as discussed previously in view of Experiment 1, 2 and 3, but in this case they were discussed with a focus on plasticity. As suggested by the SLM (Flege, 1995a, 2003), age effects are related to relevant changes and interactions between the L1 and L2 systems (Flege *et al*., 2003). L1 categories are established with increasing age (e.g. Hazan & Barett, 2000). This makes them likely to assimilate to L2 speech sounds (as PAM/PAM-L2 also suggests, Best, 1995a; Best *et al*., 2001; Best & Tyler, 2007) and also L1 experience may result in interference between the established L1 categories and the L2 categories when it comes to L2 learning as emphasized by the NLM/NLM-e (Kuhl, 1991; Kuhl *et al*., 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995; and, Kuhl *et al*., 2008) and PI account (Iverson *et al*., 2003). Based on these assumptions, it would be expected that adult L2 learners would not benefit from perceptual training methodologies, or at least not to the extent that child

L2 learners would benefit. Results from Experiment 4 showed that perceptual training was beneficial for the adult L2 learners. Even though L1 categories have already been established for the adult L2 learners, it is still possible that specialised training can aid perception of L2 speech categories by shifting attention to the 'right' dimension (e.g. Iverson & Kuhl, 1995, 1996; Goldstone, 1993, 1994; McClelland, 2001; Nosofsky, 1986). When considering improvement by adult and child groups within the same training period, children showed greater neuroplasticity to perceptual modifications. Overall, it could be assumed that the beneficial outcomes of the HVPT technique for L2 adult and child learners could be used further as a useful tool within L2 educational contexts to train a wider range of L2 contrasts (see Table 6.1).

Considering some theoretical implications of the HVPT technique, it could be suggested that in view of the current results (Experiment 4), child language learning seems to be quite different from adult language learning, at least within the context of perceptual learning. The same technique used for both child and adult L2 learners was obviously a tool that benefited both age groups, although training outcomes for L2 children were more pronounced. Another important issue is to what extent could the HVPT technique have unanticipated implications for the critical or sensitive period discussion. Findings from current as well as previous studies seem to challenge the idea of a "critical window" for language learning as well as 'stretch' previously suggested age limits and boundaries. If speech perception can be remedied through an array of perceptual training techniques such as the HVPT, then it appears that this 'window' is not so "critical" after all.

Also, the locus classicus work of Pisoni and colleagues (e.g. Pisoni *et al*., 1994) which has sparked research on perceptual training using both traditional and newer techniques (e.g. audiovisual paradigms), has highlighted the importance of high variability (i.e. the variation of stimuli and speaker that resembles more naturalistic learning contexts) as the underlying factor for L2 learning. This has been shown to also result in learners' generalization to new examples and new speakers, something that Experiment 4 in this thesis also supported. This can extend to a discussion within the context of an efficient learning method, unlike the more traditional methods of L2 instruction. The fact that the HVPT method is currently a developing approach that involves specific aspects of language learning rather than a more holistic approach to L2 learning emphasises its limitations. However, it could be proposed as a promising method that with further development could find its practical applications within the future L2 classroom.

A further implication could be realised in that the emphasis on positive outcomes of the HVPT technique (e.g. the perceptual improvement for L2 learners of a wide variety of languages such as Japanese (e.g. Bradlow *et al.*, 1997), Spanish (e.g. Aliaga-Garcia, 2011), German (e.g. Iverson & Evans, 2009), Korean (e.g. Nishi & Kewley-Port, 2008), Finnish (Ylinen *et al*., 2010), as already discussed and even long-term retention of the learnt L2 contrasts (e.g. Lively *et al*., 1994) could lead to the HVPT technique being labelled as the "magic" recipe for L2 perceptual learning. However, an observation pointed out by Iverson *et al*. (2005) is that none of the learners typically reach ceiling scores with the HVPT technique, not even with the versions using manipulated stimuli, thus the need to further improve this technique. Of course, the positive outcomes of such methodology are still emphasised particularly in view of the limited exposure of participants to such training procedures (e.g. 5-10 sessions in total) within the context of experimental research reported to date. It could be likely that further improvement may be anticipated if a HVPT technique would be applied as part of the L2 teaching curriculum.

Overall, high variability perceptual training techniques do not only have interesting pedagogical implications for L2 speech perception learning outcomes, but the findings obtained through this thesis also have a bearing on further issues in L2 speech learning, such as the ability of the adult perceptual system to remain adaptive to new input and change as learning progresses as well as children’s neuroplastic capabilities and learning capacity. Phonetic training could also be proposed as a context for assessing the role of L2 input (in terms of quality and quantity) in L2 speech learning in a controlled way.

6.4 Some observations on L2 phonological processing in tasks involving reading.

This thesis proposed a possibility of a parallel in the visual modality to what is found in the auditory modality, that is, visual or orthographic cues may also be weighted by L2 learners differently to native speakers in the perception of speech sounds of a L2. Besides L2 speech sounds being reported as assimilated or mapped on L1 speech sounds, also auditory cues being weighted differently between L1 and L2 English speakers, an additional emphasis was on the possibility that there may be orthographic cues embedded in the orthographic representations of words that could 'aid' auditory speech perception. This proposal was initially tested in Experiment 2 where such orthographic cues were 'masked' (or not available) and results confirmed initial hypotheses. It was then proposed that it could be possible that similar orthographic mappings could rule reading processing for the L2 learner, particularly when considering possible 'shared orthography' effects between the two languages (Greek – English) under consideration. The proposal was that these factors could result in confusions for the L2 learner (e.g. Oller & Ziahosseiny, 1970; Jared & Szucs, 2002; Treiman & Cassar, 1997).

The existence of orthographic cues that could affect reading performance were emphasised in chapter 5. Although this was an initial exploration in the field of reading processes, experiment 5 (chapter 5) allows for some observations on L2 reading processes by Greek learners of L2 English. A main outcome was the use of orthographic cues by Greek learners of L2 English. A sort of orthographic L1 transfer by Greek learners when reading in the L2 was observed. The existence of these orthographic cues in relation to L1 vs. L2 orthographic transparency was also emphasised. Grapheme-to-phoneme and phoneme-to-grapheme inconsistencies between the Greek and the English orthographic systems appeared to influence performance of the L2 reader, manifested by longer reaction times and error scores (also see, Glushko, 1979; Jared, McRae & Seidenberg, 1990; Lacruz & Folk, 2004; Stone, Vanhoy & van

Orden, 1997). Thus, the transparent L1 orthography (Greek) that follows one-to-one grapheme-phoneme reading rules appeared to inhibit L2 reading as English exhibits one-to-many grapheme-phoneme reading rules. Word pairs such as *weight – height*, were particularly prone to errors in phonological processing. In this way, it would appear that the native Greek speakers were particularly sensitive to non-relevant orthographic cues as compared to the native English speakers, an observation that could even propose a potential extension to the Perceptual Interference account (Iverson *et al.*, 2003). Of course, limitations in the experimental design of experiment 5 discussed in chapter 5, limit a firm claim based on current experimental outcomes. Nevertheless, this part of the thesis aims to highlight the additional facets of L2 speech learning that could be contributing to the discussion of L2 cue-weighting. It is suggested that orthographic cues could be weighted by the L2 learner in addition to perceptual weighting of acoustic cues.

Table 6.1: Summary of hypotheses, main findings, models supported and implications for L2 learning.

Experiment	Predictions / Hypotheses	Main Findings	Models specifically supported	Implications for L2 learning
1 (Chapter 2)	• Perceptual Assimilation effects. • Native language magnet effects. • Perceptual interference in relation to cue weighting. • Adult learners' experience with L2 to outperform child L2 learners.	• Perceptual assimilation. • Native language magnet effects: L1 affects L2 speech perception. • Perceptual interference in relation to cue weighting. • Adult L2 learners perform better than child L2 learners.	• PAM • SLM • NLM • PI	• L2 learners weight cues differently compared with L1 speakers. • L2 learners need to focus attention on critical than non-critical cues.
2 (Chapter 3)	• Pictures could inhibit the possible use of orthographic cues available in words. • Age-related changes.	• Pictures do inhibit the use of existing orthographic cues available in words.	• SLM • NLM • PI	• Pictorial representations could be a useful tool for L2 teaching and learning, bypassing the use of orthographic cues.
3 (Chapter 3)	• Consonantal environment could be affecting vowel perception, or provide cues for vowel identification. • Age-related changes.	• Consonantal environment does not affect vowel perception and does not affect vowel identification.	• SLM • NLM • PI	• L2 learners are more affected by vowel than consonantal word environments in CVC contexts.
4 (Chapter 4)	• Plasticity in perceptual learning for child and adult L2 learners. • HVPT technique to help L2 speech perception. • HVPT technique to benefit child and adult L2 learners.	• HVPT technique successful in training L2 learners. • Both child and adult L2 learners benefit from HVPT technique.	• PAM • SLM • NLM • PI	• HVPT technique could be used more widely in L2 education and applied to child and adult learners alike.
5 (Chapter 5)	• Observe the implications of reading processing for Greek learners of L2 English. • Greek learners could demonstrate a tendency for orthographic transfer. • Use of L1 orthographic cues at L2 reading.	• Transfer of orthographic strategies from L1 to L2. • Orthographic cues affect reading performance.	• **DRC**	• Orthography affects reading outcomes. • Training focused on drawing attention away from irrelevant orthographic cues may help L2 learners' reading processing.

6.5 Implications for L2 learning

Implications for L2 learning that emerge from this thesis (see also Table 6.1) are that L2 learners appear to weight cues differently compared with L1 speakers. Therefore, this creates a need within the L2 teaching and learning domain as to specifically lead L2 learners to focus attention on the critical rather than non-critical cues for correct perceptual identification of L2 speech sounds. The prediction that pictures could inhibit the possible use of orthographic cues available in words, or more specifically, the hypothesis that L2 learners could use L1 orthographic cues as L2 reading, was confirmed. Experiments using picture stimuli instead of words and also a purely reading experiment highlighted such tendency for L1 orthographic transfer, apart from perceptual assimilation effects found in the auditory modality.

This parallel between the auditory and visual modality is an issue that has implications for L2 learning. Teaching of an alphabetic L2 is often via grapheme-phoneme methodologies and this could often be confusing for the L2 learner, particularly when his/her L1 is a transparent orthography and the L2 is an opaque orthography or even the opposite (although this reverse scenario has not been considered or tested in this thesis). Knowledge of GPC (grapheme-to-phoneme conversion) rules can be violated, and shared orthography effects were also discussed as leading to additional reading confusions. These factors form considerable implications for L2 learning. However, it could also be suggested that taken into account they could be a beginning to revolutionizing L2 teaching and learning. One promising step forward is the HVPT technique that was tested in this thesis and proved a successful tool (further to successful previous studies discussed earlier) that could currently deal with specific aspects of L2 learning. However, the HVPT technique could be used more widely in L2 education and be applied to child and adult learners alike. Training could also be extended further by developing similar training techniques that can be focused on drawing attention away from irrelevant orthographic cues which could also help L2 learners' phonological processing and eventually L2 reading.

6.6 Limitations and future research

One major limitation of the research in this thesis is the focus in Experiments1-4 on a single L2 vowel contrast (i.e. /i:/ vs. /I/, tense – lax). This, of course, allowed for a range of tasks to be designed and performed which would not be possible if a greater range of L2 speech sound contrasts were to be tested within the time-limits and scope of this thesis. This, however, gives rise to questions such as whether findings based on this particular vowel contrast would also be true for other L2 contrasts or not. The only way of testing whether more generalized outcomes are possible would be to test more L2 contrasts (e.g. Nishi & Kewley-Port, 2007a; Lengeris, 2009).

Task difficulty could be a considerable factor for outcomes of L2 speech perception. One example is the Classification task administered as part of Experiment 1 and revealed practical difficulties for the L2 learners as well as the English native speakers alike. Task type and the degree of task difficulty should be considered very carefully at interpretation of results, especially in the case of results being concealed to chance levels. An additional example on task difficulty that is already discussed in previous chapters is the perceptual identification versus auditory discrimination task type used in most experimental paradigms of this thesis. The auditory discrimination task had consistently higher accuracy scores by L1 and L2 speakers alike, pointing to the possibility of simply being an easier task compared to requirements of the perceptual identification task. In addition, another limitation that was also discussed in Experiment 2 (Chapter 3) was the use of pictorial stimuli with L2 learners, particularly with L2 child groups. Although a number of additional control tasks (e.g. paper-based tests to examine picture vs. concept/ word meaning knowledge) were administered, an increased level for task difficulty could still pose a limitation on the discussion of experimental outcomes.

Another limitation was the potential effect of L2 proficiency level for adult participants in comparison with child participants, particularly when considering age effects and maturational constraints. This was a particularly limiting factor at discussing outcomes of experiment 1, 2, and 3 where L2 children generally had much lower scores than L2 adults. In this case, although it could be suggested that there are clear age effects given research outcomes, this suggestion was particularly limited when considering the additional difference in years of L2 education (in other words, L2 experience or difference in L2 proficiency levels) between the two groups. This limitation was reversed to some extent in Experiment 4, when L2 child participants were observed to improve more than L2 adults following HVPT intervention, despite their low starting points as well as a much lower L2 proficiency level in comparison to L2 adults. Nevertheless, it would be necessary in future experimental paradigms to better control for L2 proficiency levels between age groups in order for a discussion on age effects to be more conclusive.

Limitations of the HVPT study could be the fact that only one experimental paradigm was used, that is, training with fewer word pairs and testing with more word pairs in order to observe generalization patterns for the L2 learner. However, the use of different designs (e.g. testing more L2 contrasts rather than the trained contrast, or even using audiovisual designs (e.g. Hazan *et al*. 2005; Lim & Holt, 2011) particularly given the positive outcomes reported from previous studies, could potentially provide further insights in the context of perceptual training for the L2 population under consideration. Within the context of HVPT, one further study that might be worth exploring is to *train* Greek participants with picture stimuli. Of course, such a study would need to control for factors such as the prototypicality of the pictorial representation (see a discussion in Chapter 3). No other training study has included picture stimuli instead of word stimuli and it could therefore add new insights on what is currently known about speech perception and processing.

Two further points that could be considered as limitations for the HVPT study are the following: 1. training with natural duration stimuli and training with modified duration stimuli; this was the first study to report such training manipulation and results appeared rather promising. Training with modified duration stimuli seemed to 'shift' perception toward the right dimension (i.e. attending more to spectral rather than duration cues) which was even suggested to offer 'better' training. However, results cannot be conclusive given the limited amount of participants tested within each training category. Therefore, even though a promising manipulation within the perceptual training context, this is something that should be tested further and wider in order to be able to make more conclusive assumptions. 2. The use of the 'replay' button during the training sessions, besides its positive effects in terms of the perceptual learning process, could also be considered as a constraint when considering learning input. In other words, the option for each participant to 'replay' a given stimulus during the training sessions could support claims that there could be unequal exposure to the training material between participants of a group based on individual use of the 'replay' option. Of course, a counterargument in this case could be that the learning benefits of the 'replay' option could be considered as balancing out any effects on training input. Future perceptual training studies could perhaps also control for this factor by either restricting the use of the 'replay' option or removing it completely so as to totally equalize the amount of training input each individual participant receives.

Also, an additional limitation is observing attainment following training by an additional post-training test after the course of approximately three months as explored in some previous studies (e.g. Lively *et al*., 1994; Bradlow *et al*., 1997; Bradlow *et al*., 1999). Practical reasons would not allow such extension to be included as part of this thesis, for example time for additional travel to relevant geographic locations to test Greek participants. Given the fact that previous studies have provided evidence that training shows retention effects, it would be a nice addition in case similar data based on Greek adult and child participants could be added to the literature.

Finally, a study on reading processes could include a more powerful design, for example, additional language groups of an opaque L1 besides current groups of a transparent L1 which could be compared in their reading processes of an opaque L2. This would aid the discussion on how transparency levels of the L1 could be affecting reading in the L2. Also, the testing of child groups of various ages, although it would be imperative to control for aspects such as level of L2 reading or familiarity with the reading material used. In general, this thesis has identified the need for a better connection in the literature between speech perception and reading processes, therefore future studies should aim to further the present outcomes.

6.7 Beyond the boundaries of this study: Scope for L2 English education.

Perceptual training could be used more widely in schools since students within the European Union (EU) now have access to computers, which would make this study accessible from a practical perspective. It also requires minimal expenses to run as well as minimal instruction and supervision. A number of private language schools have already incorporated computer-learning within their curriculum, but this computer-based training so far only includes practice on grammar and vocabulary skills. Specialized high variability perceptual training which would incorporate and reflect particular learning requirements based on the learner's L1 (e.g. Greek, Spanish, Italian, and the like) and the target L2 (e.g. English vowel contrasts and more) could aim to maximize learning outcomes for EU countries and beyond. Pictures used as visual stimuli can be an exciting and attractive task type, particularly for children. Developing appropriate materials that can be made available to L2 learners (child and adult learners alike) could be a supplementary method for targeted L2 perceptual learning.

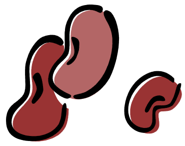
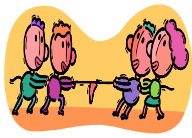
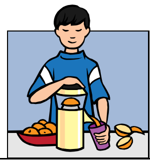
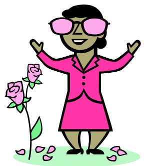
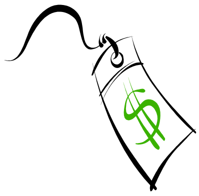
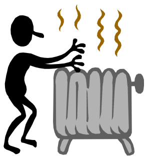
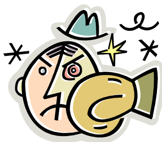
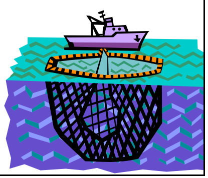
Appendix 1

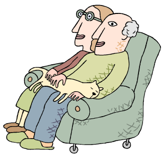
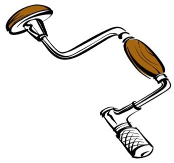
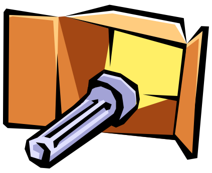
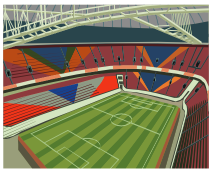
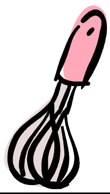
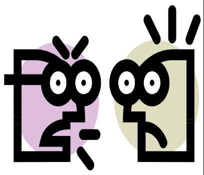
Words used in Experiment 1, arranged as minimal pairs:

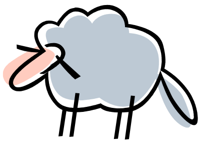
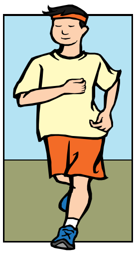
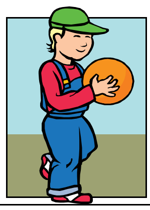
1	bean	bin
2	beat	bit
3	bead	bid
4	beaker	bicker
5	beater	bitter
6	beef	biff
7	bees	biz
8	bleep	blip
9	cheap	chip
10	deep	dip
11	dean	din
12	deed	did
13	each	itch
14	ease	is
15	feet	fit
16	feast	fist
17	fleet	flit
18	gene	gin
19	heat	hit
20	heed	hid
21	keen	kin
22	Keith	kith
23	lead	lid
24	least	list
25	leafed	lift
26	leak	lick
27	litre	litter
28	mead	mid
29	neat	knit
30	neap	nip
31	peach	pitch
32	peak	pick
33	Pete	pit
34	seat	sit
35	seed	Sid
36	seek	sick
37	seen	sin
38	sheep	ship
39	skeet	skit
40	sleep	slip
41	sleet	slit
42	sneak	snick
43	speak	spick
44	team	Tim
45	teen	tin

Appendix 2a

Picture Stimuli (and their matching words) used in Experiment 2, arranged in pairs:

1.	**bin** 	**bean** 	**2.**	**team** 	**Tim** 
3.	**feet** 	**fit** 	**4.**	**peak** 	**pick** 
5.	**teen** 	**tin** 	**6.**	**sleet** 	**slit** 
7.	**gene** 	**gin** 	**8.**	**chip** 	**cheap** 
9.	**heat** 	**hit** 	**10.**	**dip** 	**deep** 

11.	**seat** 	**sit** 	**12.**	**beat** 	**bit** 
13.	**seek** 	**sick** 	**14.**	**lid** 	**lead** 
15.	**bead** 	**bid** 	**16.**	**pitch** 	**Peach** 
17.	**fist** 	**feast** 	**18.**	**bitter** 	**Beater** 
19.	**sleep** 	**slip** 	**20.**	**beaker** 	**Bicker** 

21.	**sheep** 	**ship** 	**22.**	**litter** 	**Litre** 
23.	**lift** 	**leafed** 	**24.**	**bees** 	**biz (business)** 
25.	**pit** 	**Pete** 	**26.**	**speak** 	**Spick** 
27.	**seed** 	**Sid** 	**28.**	**knit** 	**Neat** 
29.	**leak** 	**lick** 	**30.**	**beef** 	**Biff** 

Appendix 2b

Performance for the Picture Match Task and Translation Task (1 = correct, 0 = incorrect).

Greek adult

	Picture Match										**Translation Task**									
Participant / Word	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
BEAN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BIN	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
BEAT	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BIT	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BEAD	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
BID	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1
BEAKER	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1
BICKER	1	1	0	1	0	1	1	0	1	1	0	1	1	0	1	0	1	1	1	1
BEATER	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0
BITTER	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BEEF	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
BIFF	1	0	0	1	0	1	1	0	1	0	0	1	0	1	0	0	1	1	1	1
BEES	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BIZ	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
CHEAP	1	0	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
CHIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DEEP	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DIP	1	1	0	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1
FEET	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FIT	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
FEAST	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1
FIST	1	0	0	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1
GENE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HEAT	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
HIT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LEAD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LEAFED	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	1
LIFT	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
LEAK	1	0	0	1	1	0	1	1	0	1	1	1	0	1	1	0	1	1	1	1
LICK	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1
LITRE	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

LITTER	1	0	1	0	1	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1
NEAT	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1
KNIT	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1
PEACH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PITCH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PEAK	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
PICK	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1
PETE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PIT	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1
SEAT	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
SIT	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
SEED	1	0	1	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	1	1
SID	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	1
SEEK	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SICK	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SHEEP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SHIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SLEEP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SLIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SLEET	0	0	1	1	1	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1
SLIT	1	0	1	0	1	0	1	0	0	0	1	1	0	1	1	0	1	1	0	0
SPEAK	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SPICK	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	0	1	1	0	1
TEAM	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEEN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0

Greek Child

Participant / Word	**Picture Match**										**Translation Task**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
BEAN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BIN	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
BEAT	1	1	0	0	0	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1
BIT	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1
BEAD	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
BID	1	0	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1
BEAKER	0	0	1	0	1	0	1	0	1	1	0	1	1	0	1	0	1	1	1	1
BICKER	1	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	1	1	1	1

BEATER	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0
BITTER	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1
BEEF	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
BIFF	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1
BEES	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BIZ	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1
CHEAP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CHIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DEEP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DIP	1	1	0	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1
FEET	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FIT	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
FEAST	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1
FIST	1	0	0	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1
GENE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HEAT	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
HIT	1	1	0	1	0	1	0	1	0	0	1	1	0	1	1	1	1	1	1	1
LEAD	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	0	1	1	1	1
LID	0	1	0	1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	1	0
LEAFED	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	1
LIFT	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LEAK	1	0	0	1	1	0	1	1	0	1	1	1	0	1	1	0	1	1	1	1
LICK	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1
LITRE	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1
LITTER	1	0	1	0	1	0	1	0	1	0	1	1	1	0	1	0	1	1	1	1
NEAT	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1
KNIT	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1
PEACH	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PITCH	1	1	1	0	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1
PEAK	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
PICK	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1
PETE	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1
PIT	1	0	1	1	0	1	1	0	0	1	1	1	0	1	0	1	1	1	1	1
SEAT	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
SIT	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
SEED	1	0	1	1	0	0	1	1	0	1	1	1	1	1	0	0	1	1	1	1
SID	0	1	0	1	1	1	0	1	1	0	0	1	0	1	1	0	1	1	1	1
SEEK	1	1	1	1	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1
SICK	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SHEEP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SHIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SLEEP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

SLIP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SLEET	0	1	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	1	0	1
SLIT	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	1	0	0
SPEAK	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SPICK	0	1	1	0	1	0	1	0	0	0	0	1	0	1	0	0	1	1	0	1
TEAM	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TIM	1	1	1	0	1	0	1	1	0	1	0	1	1	0	1	1	1	1	1	1
TEEN	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0

Appendix 3

Words used in the Pre- and Post-training Test of Experiment 4, arranged as minimal pairs:

1	bean	bin
2	beat	bit
3	bead	bid
4	beaker	bicker
5	beater	bitter
6	beef	biff
7	bees	biz
8	bleep	blip
9	cheap	chip
10	deep	dip
11	dean	din
12	deed	did
13	each	itch
14	ease	is
15	feet	fit
16	feast	fist
17	fleet	flit
18	gene	gin
19	heat	hit
20	heed	hid
21	keen	kin
22	Keith	kith
23	lead	lid
24	least	list
25	leafed	lift
26	leak	lick
27	litre	litter
28	mead	mid
29	neat	knit
30	neap	nip
31	peach	pitch
32	peak	pick
33	Pete	pit
34	seat	sit
35	seed	Sid
36	seek	sick
37	seen	sin
38	sheep	ship
39	skeet	skit
40	sleep	slip
41	sleet	slit
42	sneak	snick
43	speak	spick
44	team	Tim
45	teen	tin

Words used in the perceptual training sessions, arranged as minimal pairs:

1	bean	bin
2	beat	bit
3	cheap	chip
4	deep	dip
5	feet	Fit
6	gene	gin
7	heat	hit
8	lead	lid
9	least	list
10	mead	mid
11	neat	knit
12	peach	pitch
13	peak	pick
14	seat	sit
15	seek	sick
16	seen	sin
17	sleet	slit
18	team	Tim
19	teen	tin

Appendix 4a

Words and Conditions used in the Rhyme Judgment Task

		Orthographically Similar			**Orthographically Dissimilar**	
RHYME	1	plead	bead	1	hall	bawl
	2	sang	clang	2	glue	blew
	3	team	cream	3	sum	come
	4	mast	fast	4	beer	hear
	5	tight	might	5	blow	hoe
	6	hear	near	6	cram	lamb
	7	loud	proud	7	juice	loose
	8	stage	rage	8	bone	moan
	9	coot	root	9	bead	seed
	10	tall	small	10	raced	taste
	11	black	stack	11	fluff	tough
	12	land	stand	12	crane	train
	13	tare	stare	13	fled	tread
	14	plain	strain	14	sail	whale
	15	ditch	witch	15	bright	white
NON-RHYME	1	have	cave	1	hill	doll
	2	war	far	2	star	fare
	3	sew	flew	3	mice	pace
	4	boot	foot	4	trout	rough
	5	give	hive	5	fate	shoot
	6	cove	love	6	loot	shout
	7	come	nome	7	sort	shout
	8	bread	plead	8	lace	stack
	9	rough	plough	9	line	stain
	10	cost	post	10	hail	stall
	11	best	priest	11	tinge	strange
	12	pour	sour	12	cream	tram
	13	fear	swear	13	last	waist
	14	plant	want	14	crane	wine
	15	height	weight	15	belt	malt

Words used by Howard and Franklin (1996), cited in Wydell and Butterworth (1999).

Words and Conditions used in the Non-Word Rhyme Judgment Task

Rhyme Non-words			**Non-Rhyme Non-Words**		
1	yite	pight	1	trosh	desh
2	nayse	taze	2	poth	moith
3	mence	kense	3	shont	skent
4	steen	blean	4	blave	cive
5	tewt	clute	5	sturt	thart
6	trork	borque	6	trimps	blemps
7	drase	yays	7	tald	pralk
8	dryme	thime	8	trosh	moshe
9	nored	tord	9	sime	rimp
10	glinch	rinsh	10	vorve	prode
11	tude	gewd	11	bram	shamp
12	nain	tayn	12	platch	nutch
13	jide	kyde	13	flure	rire
14	gaid	vayd	14	troce	bice
15	ploit	hoight	15	bram	chrame
16	plime	jyme	16	baft	blaff
17	praik	clake	17	hesh	bish
18	trude	nood	18	trind	vond
19	daid	sade	19	wanch	tunch
20	brone	doan	20	slint	plunt
21	mobe	toab	21	filt	glid
22	soam	gome	22	bave	wace
23	staib	tabe	23	shart	prat
24	nopse	dops	24	mure	sare
25	goan	yown	25	hinth	danth

Words taken from Best (1996), cited in Wydell and Butterworth (1999).

List of Words and Conditions for Homophone Judgment Task.

Homophone Word Judgment						
	Regular			Irregular		
Match	1	blew	blue	1	ail	ale
	2	board	bored	2	bare	bear
	3	clause	claws	3	berry	bury
	4	days	daze	4	bold	bowled
	5	hair	hare	5	brake	break
	6	higher	hire	6	build	billed
	7	hole	whole	7	doe	dough
	8	hymn	him	8	earn	urn
	9	lacks	lax	9	eye	I
	10	loan	lone	10	hear	here
	11	maid	made	11	key	quay
	12	paced	paste	12	knows	nose
	13	pain	pane	13	mare	mayor
	14	pause	paws	14	moan	mown
	15	plain	plane	15	mowed	mode
	16	praise	prays	16	none	nun
	17	raise	rays	17	know	no
	18	sail	sale	18	pair	pear
	19	sighed	side	19	peace	piece
	20	sighs	size	20	seas	seize
	21	soar	sore	21	sew	so
	22	tacks	tax	22	sole	Soul
	23	tail	tale	23	some	Sum
	24	way	weigh	24	stake	Steak
	25	which	witch	25	thrown	Throne
Non-Match	1	bled	blue	1	air	Are
	2	bound	boned	2	barn	Urn
	3	clause	clams	3	bold	Boiled
	4	days	dare	4	bone	Bun
	5	hair	hard	5	built	Billet
	6	here	where	6	cry	Quay
	7	higher	hive	7	dare	Dear
	8	home	him	8	eve	I
	9	laces	lax	9	fair	Fear
	10	loan	long	10	ferry	Fury
	11	paged	paste	11	home	Hum
	12	pain	pant	12	knob	No
	13	pause	pads	13	knots	Nose
	14	plain	plant	14	mare	Major
	15	praise	prams	15	moan	moon
	16	raise	rats	16	mode	Moved
	17	sail	salt	17	new	No
	18	side	signed	18	pence	Piece
	19	signs	size	19	roe	Rough
	20	sour	soar	20	sets	Seize
	21	tail	talk	21	snake	Sneak
	22	talks	tax	22	sole	Soil
	23	wade	ward	23	stale	Steal
	24	was	weigh	24	thrown	Throng
	25	which	winch	25	wear	Were

Words used by Coltheart (1980), cited in Wydell and Butterworth (1999).

Appendix 4b

Letters of the Greek alphabet and the corresponding English letters and letter sounds.

Greek Letters	English letters *(corresponding)*	Letter sound
Α α	A a	[a]
Β β	B b	[v]
Γ γ	G g, Y y	[ɣ, ʝ]
Δ δ	D d	[ð]
Ε ε	E e	[e]
Ζ ζ	Z z	[z]
Η η	I I, E e	[i]
Θ θ	Th th	[θ]
Ι ι	I i	[i]
Κ κ	K k	[k, c]
Λ λ	L l	[l, ʎ]
Μ μ	M m	[m]
Ν ν	N n	[n]
Ξ ξ	Ks ks, X x	[ks]
Ο ο	O o	[o]
Π π	P p	[p]
Ρ ρ	R r, Rh rh	[r]
Σ σ ς	S s	[s]
Τ τ	T t	[t]
Υ υ	U u, Y y	[i]
Φ φ	Ph ph	[f]
Χ χ	Kh kh, Ch ch	[χ, ç]
Ψ ψ	Ps ps	[ps]
Ω ω	O o	[o]

Appendix 5

Brunel UNIVERSITY WEST LONDON

Linguistic / demographic background Questionnaire

Participant ID (filled by experimenter) ____________ Name: ___________

1. Date of Birth:
2. Gender:
3. Handedness: Right Left
4. Enter your native language in the space provided below. _________________
5. Enter your parent's native language below
   Mother's __________________ Father's __________________
6. What language(s) did you speak at home with either of your parents?
   Mother's _____________________Father's ________________________
7. Enter your country of origin (where you were born) in the space provided below ____
8. If you have migrated to another country since birth, enter the country in which you have resided most of your life and the age at which you migrated ____
9. At what age did you first start learning English? ___________________
10. How much of your English learning has been in language classes (in years) and in which courses? _______________________________________
11. List any English speaking communities/countries where you have lived before (where? and for how long?) ______________________________
12. How would you assess your current proficiency in reading and writing English (circle one)
   Poor Basic Intermediate Advanced

*Poor = little or no ability to read or write
*Basic = can read and write some basic words and phrases
*Intermediate = can read and write emails in English, but can have some difficulty reading complex material (e.g. Academic papers/long novels)
*Advanced = can read and write complex material (e.g. academic papers /long novels) in English

13. How would you assess your current proficiency in speaking and listening in English (circle one)

Poor Basic Intermediate Advanced

*Poor = little ability to speak or understand
*Basic = can understand and produce some basic words and phrases
*Intermediate = can understand and carry out conversations in English, but can have some difficulty with more complex discussions (e.g. academic discussions)
*Advanced = can participate in complex discussions (e.g. academic discussions) in English

14. How much time do you think you currently spend each week using English (in hours)? (Include things like listening to the radio, watching TV, films, reading in English in your estimate): ______________________________

15. Who do you regularly speak English with (select as many as appropriate)?
a. Friends b. Teachers c. Other students d. Family e.Others

16. List any other languages you speak, and give an assessment of your overall proficiency, according to the following scale (circle one)
Basic Intermediate Advanced

*Basic = can understand and produce some basic words and phrases
*Intermediate = can understand and carry out conversations in that language, but can have some difficulty with more complex discussions (e.g. academic discussions) in that language
*Advanced = Can participate in complex discussions (e.g. academic discussions) in that language

17. Did you learn any of these other languages ***before*** English and if so, at what age? ____

18. What is your most recent academic qualification (e.g. school leaving qualification, BA, other) and from which institution?

____

19. Are you studying anything at the moment? If so, which institution and course? Are you required to use English in any of these courses?

____

20. Do you currently have or ever in the past had any hearing loss or impairment? ____

21. Is your vision normal or corrected to normal? _______________________

Thank you for your help!

Bibliography

Abutalebi, J., Cappa, S.F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, 4, 179-190.

Adams, M. J. (1990). Beginning to read: Thinking and learning about print. Cambridge: MIT Press.

Ainsworth, W.A., & Millar, J.B. (1972). The effect of relative formant amplitude on the perceived identity of synthetic vowels. *Language and Speech*, 15(4), 328-341.

Aliaga-Garcia, C. (2011). Measuring perceptual cue weighting after training: A comparison of auditory vs. articulatory training methods. In: *Achievements and perspectives in SLA of speech: New Sounds 2010*,(Eds.) Magdalena Wrembel, Malgorzata Kul & Katarzyna Dziubalska-Kolaczyk. Frankfurt am Main: Peter Lang Verlag., 1, 15-28.

Allen, W. S. (1999). *Vox Graeca: A guide to pronunciation. (3rd Ed)*. Cambridge University Press.

Allport, D.A. (1979). Word recognition in reading: A tutorial review. In: P.A. Kolers, H. Bouma, & M. Wrolstad (Eds.), *Processing of Visible Language*. New York: Plenum Press.

Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22, 174-188.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition, 10*, 565-575.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. Monographs of the Society for Research in Child Development, 58 (10). 1-186.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–250.

Arvaniti, A. (1991). Rhythmic Categories: A Critical Evaluation on the Basis of Greek Data. *Proceedings of the XIIth International Congress of Phonetic Sciences*, 2, 298-301. Université de Provence, Service des Publications.

Arvaniti, A. (2000). The Phonetics of Stress in Greek. *Journal of Greek Linguistics,* 1, 9-38.

Arvaniti, A. (2007). Greek Phonetics: The State of the Art. *Journal of Greek Linguistics* (8). 97-208.

Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.). *Handbook of child psychology: cognition, perception, and language*, 2, 147–254. New York: Wiley.

Assmann, P. F., & Katz, W. F. (2005). Synthesis fidelity and vowel identification, *Journal of Acoustical Society of America*, 117, 886–895.

Badian, N. (1998). A validation of the role of preschool phonological and orthographic skills in the prediction of reading. *Journal of Learning Disabilities,* 31, 472–481.

Balaban, M. T., Snidman, N. & Kagan, J. (1997). Attention, emotion, and reactivity in infancy and early childhood. In P. J. Lang & R. F. Simons (Eds.), *Attention and orienting: Sensory and motivational processes,* 369-391. Hillsdale, NJ: Erlbaum.

Balota, D. A. & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 340-357.

Balota, D. A. & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*(1), 89-106.

Baltazani, M. (2007). Prosodic Rhythm and the Status of Vowel Reduction in Greek. *Selected Papers on Theoretical and Applied Linguistics from the 17th International Symposium on Theoretical & Applied Linguistics*, 1, 31-43. Department of Theoretical and Applied Linguistics, Thessaloniki.

Banich, M.T. & Mack, M.A. (2003). Mind, Brain, and Language: multidisciplinary perspectives. Lawrence Erlbaum Associates, Inc.

Barsalou, L.W. (1983). Ad hoc categories. *Memory & Cognition*, 11, 211-227.

Beacco, J.C. & Byram, M., (2003). Guide for the Development of Language Education Policies in Europe: From Linguistic Diversity to Plurilingual Education. Strasbourg: Language Policy Division, Council of Europe.

Beauvillain, G., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language selective access. *Journal of Memory and Language*, 26, 658-672.

Beckman, M. (1982). Segment duration and the mora in Japanese. *Phonetica*, 39, 113-35.

Bernhardt, E. B. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics*, 25, 133–150.

Besner, D. & Smith, M.C. (1992). Basic processes in reading: Is the orthographic depth hypothesis sinking? In: R. Frost & L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning*. North Holland Press.

Best, C. T. & McRoberts, G.W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46, 183–216.

Best, C. T. & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*, 305-330.

Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.). *Second language speech learning: The role of language experience in speech perception and production*. Amsterdam: John Benjamins. 13-34.

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

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

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

Best, C. T. (1995b). Learning to perceive the sound pattern of English. In C. Rovee-Collier, & L. Lipsitt (Eds.). *Advances in infancy research*. Hillsdale, NJ: Ablex. 217-304.

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

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. *Journal of Acoustical Society of America*, 109, 775–794.

Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Bion, R. A. H., Escudero, P., Rauber, A. S. & Baptista, B. O. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. In: *Proceedings of the Interspeech 2006 Conference*. 1363–1366.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706-755.

Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In: D. Birdsong (Ed.). *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates. 1-22.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*, 9-49.

Boersma, P. & Weenink, D. J. M. (2004). Praat: doing phonetics by computer (Version 4.2.17). Amsterdam: Institute of Phonetic Sciences of the University of Amsterdam. [Computer program]. Retrieved from http://www.praat.org/.

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

Bohn, O. S. & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*, 131-58.

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*. Timonium, MD: York Press. 279–304.

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

Borden, G.J., Harris, K.S., & Raphael, L.J. (2003). Speech Science Primer – Physiology, Acoustics, and Perception of Speech. 4th Edition. Lippincott Williams & Wilkins.

Borgwaldt, S.R., Hellwig, F.M., & De Groot, A.M.B. (2005). Onset entropy matters – Letter-to-phoneme mappings in seven languages. *Reading and Writing*, 18:211–229.

Bormann, T. & Weiller, C. (2012). 'Are there lexicons?' A study of lexical and semantic processing in word-meaning deafness suggests 'yes'. *Cortex*, 48,294-307.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. (2005). Mommy and Me: Familiar Names Help Launch Babies into Speech-Stream Segmentation. *Psychological Science*, 16 (4), 298 – 304.

Botinis, A. (1981). The Greek Vowels. *Working Papers, Department of Linguistics and Phonetics, Lund University. 21, 29-36.*

Botinis, A. (1989). Stress and Prosodic Structure in Greek. *Lund: Lund University Press.*

Botinis, A., Bannert, R., Fourakis, M. & Pagoni-Tetlow, S. (2002). Crosslinguistic Segmental Durations and Prosodic Typology. International Conference of Speech Prosody. Aix-en-Provence, France. ISCA Archive. 183-186.

Boyd, F. & Rozendal, M. (2003). Multicultural and Multilingual Literacy and Language: Contexts and Practices. New York: Guilford Publications, Inc.

Bradley, L. & Bryant, P. (1983). Categorizing sounds and learning to read – A causal connection. *Nature, 301*, 419-421.

Bradlow, A. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, 97, 1916-1924.

Bradlow, A.R., Akahane-Yamada, R.A., Pisoni, D.B., & Tohkura, Y. (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. (1997). Training Japanese listeners to identify English /r/and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101(10), 2299-2310.

Bradlow, A.R., Torretta, G.M., & Pisoni, D.B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20(3), 255-272.

Breen, A., Bowers, E., & Welsh, W. (1996). An Investigation into the Generation of Mouth Shapes for a Talking Head. *Proceedings of ICSLP,* 96 (4). 2159-2162.

Bridgeman, B. (1987). Is the dual route theory possible in the phonetically regular languages?. *Behavioral and Brain Sciences*, 10, 331-332.

Brizic, K. (2006). The secret life of languages: Origin-specific differences in L1/L2 acquisition by immigrant children. *International Journal of Applied Linguistics,* 16 (3), 339–362.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar position. *Journal of the Acoustical Society of America*, 117, 3890–3901.

Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14, 136-193.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory*. Malden, MA: Blackwell. 6-63.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28,* 874-886.

Bryant, P. (1986). Phonological skills and learning to read and write. In: B. R. Foorman & A. W. Siegel (Eds.). *Acquisition of reading skills: Cultural constraints and cognitive universals*. Hillside: Lawrence Erlbaum.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, Pussycat? On talking to babies and animals. *Science*, 296, 1435.

Byram, M., Gribkova, B., & Starkey, H. (2002). Developing the Intercultural Dimension in Language Teaching: A Practical Introduction for Teachers. Strasbourg: Languages Policy Division, Council of Europe.

Carlson, R., Granström, B., & Klatt, D. (1979). Vowel perception: the relative perceptual salience of selected acoustic manipulations. *Quarterly Progress Status Report*, Stockholm, Sweden, No. 3-4, 73-83.

Carney, A. E., Widin, G. P. & Viemeister, N.F. (1977). Noncategorical perception of stop consonants differing in VOT. *The Journal of the Acoustical Society of America*, 62, 961-970.

Carter, C. S., Mintun, M. & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An H2150 PET study of Stroop task performance. *NeuroImage*, 2, 264-272.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, 34, 371-387.

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

Cenoz J. & Jessner, U. (Eds.) (2000). *English in Europe: The acquisition of a third language*. Clevedon: Multilingual Matters.

Cerviño, E. & Mora, J.C. (2009). Duration as a phonetic cue in the categorisation of */iː/-/ɪ/* and */s/-/z/* by Spanish/Catalan learners of English. *Journal of the Acoustical Society of America*, 125 (4), 2, 2764.

Chiba, T., & Kajiyama, M. (1941). The Vowel: Its Nature and Structure. Tokyo Publishing Co, Tokyo, 174-226.

Chiswick, B. R., & Miller, P. W. (2008). A test of the critical period hypothesis for language learning. *Journal of Multilingual and Multicultural Development*, 29, 16-29.

Cole, R., & Jakimik, J. (1980). A model of Speech Perception. Hillsdale, N.J: Lawrence Earlbaum Associates.

Coltheart M. (2004). Are there lexicons?. *Quarterly Journal of Experimental Psychology*, 57 A(7): 1153e1171.

Coltheart M. (2006). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology*, 23(1): 96e107.

Coltheart, M. (1980). Reading phonological recoding, and deep dyslexia. In Coltheart, M., Patterson, K., and Marshall, J.C. (Eds.), *Deep Dyslexia*. London: Routledge and Kegan Paul.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance,* 20(6), 1197-1211.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review,* 108 (I), 204-256.

Comeau, L., Cormier, P., Grandmaison, E. & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 91(1), 22–43.

Content, A. & Peereman, R. (1992). Single and multiple process models of print to speech conversion. In J. Alegria, D. Holender, J. Morais, & M. Raft (Eds.), *Analytic approaches to human cognition*. Amsterdam: Elsevier.

Cornelissen, P., Munro, N., Fowler S., & Stein J. (1993). The stability of binocular fixation during reading in adults and children. *Developmental Medicine & Child Neurology*, 35, 777–787.

Cossu, G., Gugliotta, M., & Marshall, J. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes?. *Reading and Writing*, 7(1). 9-22.

Cranshaw, A. (1997). A study of Anglophone native and near-native linguistic and metalinguistic performance. Unpublished doctoral dissertation. Universite de Montreal.

Crystal, D. (1991). The Cambridge Encyclopedia of Language. CUP.

Cummins, J. (1991). Language learning and bilingualism. Tokyo: Sophia University. In: Curtain, H. & Pescola, C. A. B., (1994). *Languages and Children: Making the Match*. New York: Longman.

Curtiss, S. R. (1977). Genie: A linguistic Study of a Modern day 'Wild Child'. Academic Press: New York.

Curtiss, S. R. (1980). The critical period and feral children. *UCLA Working Papers in Cognitive Linguistics*, 2, 21-36.

Curtiss, S.R. (1988). Abnormal language acquisition and the modularity of language. In: F. Newmeyer, vol. 2 (Ed.). *Linguistics: The Cambridge Survey*, Cambridge University Press, Cambridge.

Cutler, A. & Broersma, M. (2005). Phonetic precision in listening. In: W. J. Hardcastle, & J. Mackenzie Beck (Eds.). A figure of speech: A Festschrift for John Laver. London: Lawrence Earlbaum. 63-91.

D'Angiulli, A., Siegel, L. S. & Serra, E. (2001). The development of reading in English and Italian in bilingual children. *Applied Psycholinguistics*, 22, 479–507.

Dalston, R. M. (1975). Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America*, 57, 462–469.

de Groot, A. M. B., Borgwaldt, S., Bos, M. & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language,* 47(1), 91-124.

de Groot, A. M. B., Delmaar, P., & Lupker, S. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, 53A, 397-428.

de Jong, K. J. & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30, 53–75.

de Jong, K. J. (2004). Stress, lexical focus and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics,* 32, 493-516.

de Jong, P. F. & Van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. Scientific Studies of Reading, 6(1), 51–77.

Denton, C. A., Hasbrouk, J. E., Weaver, L. R. & Riccio, C. A. (2000). What do we know about phonological awareness in Spanish? *Reading Psychology*, 21, 335–352.

Diaz, B., Baus, C., Escera, C., Costa, A. & Sebastian-Galles, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences*, 105(42), 16083–16088.

Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (1999). Recognizing cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41, 496-518.

Dijkstra, T., De Bruijn, E., Schriefers, H., & Ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism, Language and Cognition*, 3, 55-64.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism, Language and Cognition*, 1, 51-66.

Donnenwerth-Nolan, S., Tanenhaus, M.K. & Seidenberg, M.S. (1981). Multiple code activation in word recognition: evidence from rhyme monitoring. *Journal of Experimental Psychology & Human Learning*, 7, 170 –180.

Donovan, R. (1996). *Trainable Speech Synthesis*. PhD Thesis, Cambridge University Engineering Department, England.

Durgunoglou, A. Y. & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading & Writing*, 11, 281-299.

Edwards, J. (1994). Multilingualism. London: Routledge.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.

Ellis, N. C. & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics*, 22, 571-599.

Ellis, N. C. (2002). Frequency effects in language processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. SSLA 24, Cambridge University Press. 143-188.

Elman, J.L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48, 71-99.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). Rethinking innateness: A connectionist perspective on development. Cambridge, MA: MIT Press.

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

Escudero, P. (2000). Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels in Spanish speakers. Unpublished MSc. thesis, University of Edinburgh.

Escudero, P. (2005). Linguistic perception and second language acquisition. Explaining the attainment of optimal phonological categorization. Unpublished PhD thesis, Utrecht University, The Netherlands.

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., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36. 345–360.

Fant, G. (1973). Speech sounds and features. Cambridge, MA: MIT Press.

Fernald, A., Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10 (3), 279–293.

Flege, J. (1988). The production and perception of speech sounds in a foreign languages. In: H. Winitz (Ed). *Human Communication and Its Disorders, A Review*. Norwood, N.J.: Ablex, 224-401.

Flege, J. E. & Hillenbrand, J. (1986). Differential use of temporal cues to the /s/- /z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America*, 79, 508-517.

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

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

Flege, J. E. & Murno, M.J. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, 95 (6), 3623-3641.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15*, 443-455.

Flege, J. E. (1987). The production of "new" and "similar" phonemes in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47-65.

Flege, J. E. (1991a). Perception and production: The relevance of phonetic input to L2 phonological learning. In: T. Heubner, & C. Ferguson (Eds.). *Crosscurrents in second language acquisition and linguistic theory*. Philadelphia: John Benjamins. 249-284.

Flege, J. E. (1991b). The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology,* 43, 701-731.

Flege, J. E. (1992). Speech learning in a second language. In: C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.). *Phonological development: Models, research, and application*. Timonium, MD: York Press. 565-604.

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

Flege, J. E. (1995b). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics*, 16, 425-442.

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

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In: A. Meyer & N. Schiller (Eds.). *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities,* Berlin: Mouton de Gruyter, Berlin. 319-355.

Flege, J. E., & Hammond, R. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition, 5*, 1-17.

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

Flege, J. E., Murray J. M. & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, *97*, 3125–3134.

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. & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78-104.

Flege, J., 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, 153-175.

Foorman, B. R. (1994). The Relevance of a connectionist model of reading for 'The Great Debate'. *Educational Psychology Review*, 6 (1), Plenum Publishing Corporation. 25-47.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635.

Fourakis, M. (1991*). Tempo, Stress, and vowel reduction in American English. Journal of Acoustical Society of America*, 90, 1816-1827.

Fourakis, M., Botinis, A., & Katsaiti, M. (1999). Acoustic Characteristics of Vowels. *Phonetica,* 56, 28-43.

Fowler CA, Sramko V, Ostry DJ, Rowland SA, Hallé PA. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*, 36, 649–663.

Fox, R. A., Flege, J. E. & Murno, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of Acoustical Society of America*, 97 (4), 2540-2551.

Francis, A .L. & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 349–366.

Frederiksen, J. R. & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 361-379.

Frith, U., Wimmer, H. & Landerl, K. (1998). Differences in phonological recoding in German and English speaking children. *Scientific Studies of Reading*, 2, 31-54.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology*, 74(2), 159-180.

Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (Eds.). (1998). Cognitive neuroscience, the biology of the mind. New York, London: Norton &Co.

Geeraerts, D. (1986). On necessary and sufficient conditions. *Journal of Semantics*, 5 (4), 275-291.

Gervain, J. & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, 61, 191–218.

Giannakopoulou, A. (2003). Prototype Theory: An Evaluation, in *ECLOGA Working Papers*, 3, On-line journal: http://www.strath.ac.uk/ecloga/ contents3.html.

Giannakopoulou, A. (2004). Passive Voice in Translation of English Into Greek: a Case Study Based on 9/11 by Chomsky, N. (2001) and its Greek Translation, In: *Working Papers in English and Applied Linguistics*, 10, Cambridge University Press. 103-153.

Giannakopoulou, A., Uther, M., & Ylinen, S. (2011). Phonetic cue-weighting in the acquisition of a second language (L2): evidence from Greek speakers of English. In: Achievements and perspectives in SLA of speech: New Sounds 2010. In: Magdalena Wrembel, Malgorzata Kul and Katarzyna Dziubalska-Kolaczyk. (Eds.). Frankfurt am Main: Peter Lang Verlag, (1), 91-102.

Giegerich, H. J. (2001). English Phonology –An Introduction. Cambridge University Press.

Gimson, A. C. (1962). An Introduction to the Pronunciation of English. London: Arnold. [6th edn. (2001), edited by A. Cruttenden.].

Girelli, L., Lucangeli, D. & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology,* 76(2), 104-122.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*(4), 674-691.

Goldin-Meadow, S. (1982). The resilience of recursion: A study of a communication system developed without a conventional model. In: Language Acquisition: State of the Art. (Eds.). E. Wanner & L.R. Gleitman. Cambridge University Press. 51-77.

Goldstone, R. (1993). Feature distribution and biased estimation of visual displays. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 564–579.

Goldstone, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General,* 123, 178–200.

Golestani, N. & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*, 109, 55–67.

Goswami, U. (2000). Phonological Representations, Reading Development and Dyslexia: Towards a Cross-Linguistic Theoretical Framework. *Dyslexia,* 6, 133-151.

Goswami, U. (2008). Cognitive Development: The Learning Brain. Psychology Press.

Goswami, U., Porpodas. C. & Wheelwright, S. (1997). Children's orthographic representation in English and Greek. *European Journal of Psychology of Education,* 12, 273-292.

Goswami, U., Ziegler, J., Dalton, L. & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units?. *Applied Psycholinguistics*, 24, 235-247.

Goswami, U., Ziegler, J.C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, 92, 345-365.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia*, 9(3), Pergamon Press. 317-323.

Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics,* 12, 91–114.

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (1999). Reading homographs: Orthographic, phonologic and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 561-574.

Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology,* 25, 577-588.

Grosjean, F., Ping, L., Munte, T.F. & Rodriguez-Fornells, A. (2003). Imaging bilinguals: when the neurosciences meet the language sciences. *Bilingualism: Language and Cognition*, 6, 159-165.

Guion, S. G., Flege, J. E., Akahane-Yamada R. & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America,* 107, 2711-2724.

Hall, C. (2003). Modern German pronunciation: An introduction for speakers of English (2nd ed.). New York: Manchester University Press.

Hallé, P. A., Chéreau, C., & Segui, J. (2000).Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, *43*, 618–639.

Hardison, D. (2003). Acquisition of second-language speech: effects of visual cues, context and talker variability. *Applied Psycholinguistics,* 24, 495-522.

Hattori, K. & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. Journal Acoustical Society of America, 125, 469–479.

Hattori, K., Iverson, P. (2007). Examination of similarity between English /r/, /l/, and Japanese flap: An investigation of best exemplars by English and Japanese speakers. *16th International Congress of Phonetic Sciences*. Saarbrücken, Germany. 1777-1780.

Hawkins, S. & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*. 35(2). 183-199.

Hawks, J. W. & Fourakis M. S. (1995). The perceptual vowel spaces of American English and modern Greek: a comparison. *Language and Speech*, 38, 237-252.

Hazan, V. & Barrett, S. (2000). The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics*, 28, 377-396.

Hazan, V., Sennema, A., Faulkner, A. (2004). Effect of intensive audiovisual perceptual training on the perception and production of the /l/-/r/ contrast for Japanese learners of English. *Proceedings of International Conference for Spoken Language Processing, Jeju, October 2004*. Vol. 2. 1145-1148.

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. *Journal of the Acoustical Society of America*, 119, 1740–1751.

Hazan, V., Sennema, A., Iba, M. & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants in Japanese learners of English. *Speech Communication*, 47 (3), 360-378.

Herschensohn, J. R. (2007). Language development and age. Cambridge University Press.

Hertz-Pannier, L., Chiron, C., Jambeque, I., Renaux-Kieffer, V., Van de Moortele,P.R., Delalande, O., Fohlen, M., Brunelle, F., & Le Bihan, D. (2002). Late plasticity for language in a child's non-dominant hemisphere: A pre- and post-surgery fMRI study. *Brain*, 125, 361-372.

Hillenbrand, J. M., & Nearey, T. M. (1999). Identification of resynthesized /hVd/ utterances: Effects of formant contour. *Journal of the Acoustical Society of America*, 105 (6), 3509-3523.

Hillenbrand, J. M., Clark, M. J. & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America*, 108, 3013–3022.

Hillenbrand, J., Clark, M., & Nearey, T. (2001). Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, 109(2), 748–763.

Hillenbrand, J., Getty, L.A., Clark, M.J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 686-713.

Hirata, Y. (2004) .Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *Journal of the Acoustical Society of America*, 116, 2384-2394.

Hirata, Y., Whitehurst, E. & Cullings, E. (2007). Training native English speakers to identify Japanese vowel length contrast with sentences at varied speaking rates. *Journal of the Acoustical Society of America*, 121, 3837-3845.

Hoff-Ginsburg, E. (1997). Language Development. Pacific Grove, CA: Brooks-Cole.

Hojen, A. & Flege, J. (2006). Early learners' discrimination of second-language (L2) vowels. *Journal of the Acoustical Society of America*. 119, 3072-3084.

Holst, T. & Nolan, F. (1995). The Influence of Syntactic Structure on [s] to [ʃ] Assimilation. *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV,* Bruce Connell & Amalia Arvaniti (Eds.).Cambridge University Press. 315-333.

Holt, L. L. & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119, 3059-3071.

Howard, D. & Franklin, S. (1996). Missing the meaning? MIT Press, Cambridge, M.A.

Hulme, C., Hatcher, P. J., Nation, K., Brown, A., Adams, J. & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology*, 82, 2–28.

Hunt, R. R. & Ellis, H. C. (2004). Fundamentals of Cognitive Psychology (7th Edition). McGraw-Hill Companies, Inc.

Hyltenstam, K. & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 539-588). Oxford: Blackwell.

Iverson, P. & Evans, B. G. (2007a). Auditory training of English vowels for first-language speakers of Spanish and German. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany, 1625-1628.

Iverson, P. & Evans, B. G. (2007b). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement and duration. *Journal of the Acoustical Society of America*, 122, 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. *Journal of the Acoustical Society of America*, 126(2), 866-877.

Iverson, P. & Kuhl, P. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America, 99,* 1130–1140.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, 97, 553-562.

Iverson, P., Ekanayake, D., Hamann, S., Sennema, A. & Evans, B. G. (2008). Category and perceptual interference in second-language phoneme learning: An examination of English /w/-/v/ learning by Sinhala, German, and Dutch speakers. *Journal of Experimental Psychology of Human Perception and Performance*, 34, 1305–1316.

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. *Journal of the Acoustical Society of America*, 118, 3267-3278.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47–B57.

Iverson, P., Smith, C. A. & Evans, B. G. (2006). Vowel recognition via cochlear implants and noise vocoders: Effects of formant movement. *Journal of the Acoustical Society of America*, 120, 3998–4006.

Jacewicz, E. & Fox, R. A. (2008). Amplitude variations in co-articulated vowels. *Journal of the Acoustical Society of America*, 123 (5). 2750–2768.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception and Psychophysics*, 40, 205-215.

Jamieson, D.G., & Morosan, D.E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, 43, 88-96.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words?. *Journal of Memory and Language*, 44, 2-31.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph recognition. *Bilingualism: Language and Cognition*, 5(3), 225-239.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *The Journal of Memory and Language*, *29*, 687–715.

Jessner, U. (2006). Linguistic awareness in multilinguals: English as a third language. Edinburgh: Edinburgh University Press.

Jessner, U. (2008). Teaching third languages: Findings, trends and challenges. *Language Teaching*, 41(1), 15–56. Cambridge University Press.

Johnson, E. K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548-567.

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, 60-99.

Jongman, A., Fourakis, M., & Sereno, J. A. (1989). The acoustic vowel space of Modern Greek and German. Language and Speech, 32, 221-248.

Joseph, B.D. & Philippaki-Warburton, I. (1987). *Modern Greek*. Croom Helm. London.

Juphard, A., Carbonnel, S. & Valdois, S. (2004). Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition*, 55(2), 332-340.

Jusczyk, P. (1989). Developing phonological categories from the speech signal. In: *Proceedings of the International conference on Phonological development*, Stanford University.

Jusczyk, P. W. & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.

Jusczyk, P. W., Cutler, A., & Redanz, L. (1993a). Infants' sensitivity to predominant stress patterns in English. *Child Development*, 64, 675–687.

Jusczyk, P. W., Friederici, A. D, Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402–420.

Jusczyk, P., Hohne, E. & Bauman, A. (1999a). Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics*, 61, 1465–1476.

Jusczyk, P., Houston, D. & Newsome, M. (1999b). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159–207.

Katsura, A., Flege, J.E., Guion, S.G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–250.

Katz, L., & Feldman, L. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 157-166.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning*. North Holland Press.

Keenan, T. (2002). An Introduction to Child Development. Sage Publications Ltd.

Kemps, R. J. J. K., Ernestus, M., Schreuder, R., & Baayen, R.H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory and Cognition*, 33 (3), 430-446.

Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S. Swansburg, M.L., Lee, K., Xie, X., Huang, H., Ye, H.H., Zhang, K., & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior & Development*, 32, 59–71.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5), 1208–1221.

Knudsen, E.I. (1999). Early experience and critical periods. In: M.J. Zigmond (Ed.), Fundamental Neuroscience. San Diego, CA: Academic Press. 637-654.

Knudsen, E.I. (2004). Sensitive periods in the development of the brain and behavior. Journal of Cognitive Neuroscience, 16, 1412–1425.

Kontosopoulos, N., Ksiromeritis, N. & Tsitsa, A. (1988). Akustikes erevnes epi ton elinikon fthogon (Acoustic research on the Greek sounds). *Glossologia.* 5(6). 103-122.

Koutsoudas, A. & Koutsoudas, O. (1962). A contrastive analysis of the segmental phonemes of Greek and English. *Language Learning*, 12, 211-230.

Krashen, S. D. (1993). The power of reading. Englewood, CO: Libraries Unlimited.

Kroll, J. F. & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149-174.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrica*, 29, 1–27.

Kucer, S. D. (2005). Dimensions of literacy (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the “perceptual magnet effect”. In: W. Strange (Ed.), *Speech perception and linguistic* *experience: Issues in cross-language research*. Baltimore: York Press. 121-154.

Kuhl, P. K. (1992). Infants' perception and representation of speech: Development of a new theory. In: *Proceedings of the International Conference on Spoken Language Processing*. J. J. Ohala, T.M. Nearey, B. L. Derwing, M.M. Hodge, & G.E. Wiebe (Eds.). University of Alberta, Edmonton, Alberta, 449-456.

Kuhl, P. K. (1993c). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics*, *21*, 125–139.

Kuhl, P. K. (1993b). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In: *Developmental Neurocognition Speech and Face Processing in the First Year of Life*, B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. Mac-Neillage. & J. Morton (Eds.). Kluwer Academic, Boston. 250-274.

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, 4, 812–822.

Kuhl, P. K. (1998). The development of speech and language. In: T. C. Carew, R. Menzel & C. J. Shatz (Eds.). *Mechanistic relationships between development and learning.* New York: Wiley. 53–73.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(*22*), 11850–11857.

Kuhl, P. K. (2004). Early Language Acquisition: Cracking the Speech Code. *Nature Reviews, Neuroscience*, 5, 831-843.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). Early phonetic perception as a gateway to language: New data and Native Language Magnet Theory, expanded (NLM-e). *Philosophical Transactions of the Royal Society*, 363(B), 979-1000.

Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. (2003). Foreign language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Procedures of the National Academy of Sciences USA, 100, 9096–9101.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608.

Kuhl, P.K. (1991). Human adults and human infants show a ‘perceptual magnet effect’ for the prototypes of speech categories, monkeys do not’. Perception & Psychophysics, 50 (2), 93-107.

Kuhl, P.K. (1993a). Infant speech perception: A window on psycholinguistic development. *International Journal of Psycholinguistics*. 33-56.

Kuhl, P.K. (2010). Brain Mechanisms in Early Language Acquisition. *Neuron*, 67, 713-727.

Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277 (5326), 684–686.

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, 13–21.

Lacruz, I., & Folk, J. R. (2004). Feedforward and feedback consistency effects for high- and low frequency words in lexical decision and naming. *The Quarterly Journal of Experimental Psychology*, 57(A), 1261–1284.

Ladefoged, P. & Maddieson, I. (1996). The Sounds of the World's Languages. Oxford: Blackwell.

Ladefoged, P. (2001). A course in phonetics (4th Ed.). Orlando, FL: Harcourt College Publishers.

Ladefoged, P. (2005). Vowels and consonants: An introduction to the sounds of languages (2nd ed.). Malden, MA: Blackwell Publishing.

Lambacher, S., Martens, W., & Kakehi, K. (2002). The influence of identification training on identification and production of the American English mid and low vowels by native talkers of Japanese. In: *Procedures of the 7th International Conference on Spoken Language Processing*, 245–248.

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, 227-247.

Landerl, K., Frith, U., & Wimmer, H. (1996). Intrusion of orthographic knowledge on phoneme awareness: Strong in normal readers, weak in dyslexic readers. *Applied Psycholinguistics*, *17*, 1–14.

Landerl, K., Wimmer, H. & Frith, U. (1997). The impact of orthographic consistency on dyslexia: a German-English comparison. *Cognition*, 63, 315–334.

Lane, H. (1976). The Wild Boy of Aveyron: A history of the education of retarded, deaf and hearing children. Cambridge, MA: Harvard University Press.

Laurence, S. & Margolis, E. (Eds.) (1999). Concepts and Cognitive Science. London: Blackwell.

Lee, C.Y., Huang, H.W., Kuo, W.J., Tsai, J.L., & Ovid Tzeng, J.L. (2009). Cognitive and neural basis of the consistency and lexicality effects in reading Chinese. *Journal of Neurolinguistics*. 1-18.

Lee, J., Perrachione, T., Dees, T. & Wong, P. C. M. (2007). Differential effects of stimulus variability and learners' pre-existing pitch perception ability in lexical tone learning by native English speakers, In: *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrucken, Germany, 1589–1592.

Lehrer, A. (1989). Prototype theory and its implications for lexical analysis. In: Tsohatzidis, S.L. (ed.), *Meanings and Prototypes: Studies in Linguistic Categorization.* London: Routledge.

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19, 585-611.

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. Journal of the Acoustical Society of America, 128 (6), 3757-3768.

Lengeris, A. (2008). The effectiveness of auditory phonetic training on Greek native speakers' perception and production of Southern British English vowels. *Proceedings of the 2nd ISCA Workshop on Experimental Linguistics*, ExLing 2008, Athens, Greece.

Lengeris, A. (2009). Perceptual Assimilation and L2 Learning: Evidence from the Perception of Southern British English Vowels by Native Speakers of Greek and Japanese. *Phonetica*, 66, 169-187.

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, Saarbrueken, Germany. Saarbrueken, Germany. 1641-1644.

Lenneberg, E. H. (1967). Biological foundations of Language. New York: Wiley.

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.

Liberman, A. M., Harris, K. S., Hoffman, H. & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368.

Liberman, A. M., Harris, K., Eimas, P., Lisker, L., & Bastian, J. (1961a). An effect of learning on speech perception: the discrimination of durations of silence with and without phonemic significance. *Language Speech*, 4, 175-95.

Liberman, A.M., Cooper, F.S., Shankweiler, D.S. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.

Liberman, A.M., Harris, K.S., Kinney, J.A., & Lane, H. (1961b). The discrimination of relative onset time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology*, 61, 379-88.

Liberman, I., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Speech Education*, 6, 8-17.

Lim, S. J. & Holt, L. L. (2011). Learning Foreign Sounds in an Alien World: Videogame Training Improves Non-native Speech Categorization. Cognitive Science, 35, 1390-1405.

Lintunen, P. (2004). Pronunciation and phonemic transcription: A study of advanced Finnish learners of English (Anglicana Turkuensia No. 24). Turku, Finland: University of Turku.

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. *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. *Journal of the Acoustical Society of America*, 96, 2076–2087.

Logan, J. D., 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, 874–886.

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: Theoretical and methodological issues*, Timonium, MD: York Press. 351-378.

Loizidou-Ieridou, N., Masterson, J., & Hanley, J.R. (2010). Spelling development in 6–11-year-old Greek-speaking Cypriot children. *Journal of Research in Reading*, 33(3), 247–262.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.

Lotto, A. J., Sato, M. & Diehl, R. (2004). Mapping the task for the second language learner: the case of Japanese acquisition of /r/ and /l/. In: From sound to sense. J. Slitka, S.Manuel & M. Matthies (Eds.). Cambridge, MA: MIT Press.

Lukatela, G., Popadic, D., Ognjenovic, P. & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory and Cognition, 8*, 415-423.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics*, 46, 187–200.

MacKay, I., Flege, J., Piske, & Schirru, C. (2001). Category restructuring during second-language (L2) speech acquisition. *Journal of the Acoustical Society of America*, 110, 516-528.

MacKay, I., Meador, D., & Flege, J. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58, 103-125.

Mackridge, P. (1985). The Modern Greek language: a descriptive analysis of Standard Modern Greek. Oxford University Press.

Magnuson, J.S., Yamada, R.A., Tohkura, Y., Pisoni, D.B., Lively, S.E. & Bradlow, A.R. (1995). The role of talker variability in non-native phoneme training. *Acoustical Society of Japan*. 393-394.

Major, R. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Erlbaum.

Major, R. C. (2002). The ontogeny and phylogeny of second language phonology. In J. Leather & A. James (eds.), *New Sounds 2000: Proceedings of the 4th International Symposium on the Acquisition of Second-Language Speech*. Klagenfurt: University of Klagenfurt. 223–235.

Manis, F. R., Custodio, R. & Szeszulski, P. (1993). Development of phonological and orthographic skill: A 2-year longitudinal study of dyslexic children. *Journal of Experimental Child Psychology, 56*, 64-86.

Marinova-Todd, S. H. (2003). Comprehensive analysis of ultimate attainment in adult second language acquisition. Unpublished doctoral dissertation, Harvard University.

Marshall, J. C. & Newcombe, F. (1973). Patterns of Paralexia: A Psycholinguistic Approach. *Journal of Psycholinguistic Research*, Plenum Publishing Corporation, N.Y., 2(3), 175-199.

Mattys, S. & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91–121.

Mattys, S., Jusczyk, P., Luce, P., & Morgan, J. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465–494.

Mayberry, R. (1993). First-language acquisition after childhood differs from second-language acquisition: The case of American Sign Language. *Journal of Speech and Hearing Research*, 36, 1258-1270.

Mayberry, S., Fischer, S., & Hatfield, N. (1983). Sentence repetition in American Sign Language. In: Language in Sign: International Perspectives on Sign Language. J. Kyle & B. Woll (Eds.), London: Croom Helm. 83-97.

Maye, J., Werker, J.F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(B). 101-111.

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, 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: predictions of a hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience*, 2, 89–108.

McClelland, J. L. (2001). Failures to learn and their remediation: A Hebbian account. In J. L. McClelland & R. S. Siegler (Eds.), *Mechanisms of cognitive development,* Mahwah, NJ: Erlbaum. 97–121

McClelland, J., Fiez, J., & McCandliss, B. (2002). Teaching the /r/-/l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology & Behavior*, 77, 657-662.

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748.

Miozzo M & Gordon P. (2005). Facts, events, and inflection: When language and memory dissociate. *Journal of Cognitive Neuroscience,* 17(7): 1074e1086.

Miozzo M. (2003).On the processing of regular and irregular forms of nouns and verbs: Evidence from neuropsychology. *Cognition*, 87(2): 101e127.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics*, 18, 331–340.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H. & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 18(3), 452-467.

Moon, S. J. & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of Acoustical Society of America*. 96. 4-55.

Mora, J. C. & Fullana, N. (2007). Production and perception of English /iː/ - /ɪ/ and /æ/ - /ʌ/ in a formal setting: Investigating the effects of experience and starting age. *Proceedings of the 16th International Congress of the Phonetics Sciences*, Saarbrucken, Germany, 207-221.

Morgan, J. L., & Saffran, J, R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911-936.

Morgan, J. L., Bonamo, K. & Travis, L. (1995). Negative evidence on negative evidence. *Developmental Psychology*, 31, 180–197.

Morosan, D.E. & Jamieson, D.G. (1989). Evaluation of a technique for training new speech contrasts: generalization across voices, but not word-position or task. *Journal of Speech and Hearing Research*, 32(3), 501-11.

Morrison, G.S. (2002). Effects of L1 duration experience on Japanese and Spanish listeners' perception of English high front vowels. Unpublished master's thesis. Simon Fraser University, Burnaby, British Columbia, Canada.

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, 58 (3), 351-362.

Munro, M. J. (1993). Production of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, 36, 39-66.

Nakai, S. (1997). Perceptual factors in the acquisition of FL vowels: The locations of vowel prototypes. *Proceedings of the 1997 Edinburgh Linguistics Department Postgraduate Conference.* Available Online: http://www.ling.ed.ac.uk/~pgc.

Nakai, S. (1999). The perceptual distance between vowels: Effects of prototypicality and extremity. Unpublished doctoral dissertation, Edinburgh University.

Nakatani, L. H., O'Connor, K. D. & Aston C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38, 84-106.

Navarro, T. (1968). Studies in Spanish phonology. Coral Gables, FL: University of Miami Press.

Neville, H., Coffey-Corina, S., Lawson, D., Fischer, A., Emmorey, K., & Bellugi, U. (1997). Neural systems mediating American Sign Language: Effects of sensory experience and age of acquisition. *Brain and Language*, 57, 285-308.

Neville, H., Mills, D., & Lawson, D. (1992). Fractionating language: Different neural sub-systems with different sensitive periods. *Cerebral Cortex*, 2, 244-258.

Newport, E. & Supalla, T. (1992). Critical period effects in the acquisition of a primary language: The influence of maturational state on the acquisition of complex morphology in American Sign Language. University of Rochester.

Newport, E. (1984). Constraints on learning: Studies in the acquisition of American Sign Language. *Papers and Reports on Child Language Development*, 23, 1-22.

Newport, E.L., & Aslin, R.N. (2000). Innately constrained learning: Blending old and new approaches to language acquisition. In S.C. Howell, S.A. Fish, and T. Keith-Lucas (Eds.), Proceedings of the 24th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.

Nicolaidis, K. & Rispoli, R. (2005). The Effect of Noise on Speech Production: An Acoustic Study. *Studies in Greek Linguistics,* 25, 415-426.

Nicolaidis, K. & Sfakianaki, A. (2007). An acoustic analysis of vowels produced by Greek speakers with hearing impairment. *Proceedings of the ICPHS 2007*. Saarbruecken. Germany. (6-10 August 2007).

Nicolaidis, K. (1991). Coarticulatory Strategies in Greek and English VCV sequences. *Proceedings of the 5th Symposium on the Description and/or Comparison of English and Greek*. Thessaloniki: Aristotle University, School of English. 113-140.

Nicolaidis, K. (2003). Acoustic Variability of Vowels in Greek Spontaneous Speech. *Proceedings of the XVth International Congress of Phonetic* Sciences, Universidad Autónoma de Barcelona. 3221-3224.

Nikopoloulos, D., Goulandris, N., & Snowling, M.J. (2003). Developmental dyslexia in Greek. In: *Dyslexia in different languages – Cross linguistic comparisons*. Goulandris, N. (Ed.). Athenaum Press Limited, Gateshead, Tyne & Wear. 53-67.

Nishi, K. & Kewley-Port, D. (2007a). Second language vowel perception training: Effects of set size, training order, and native language. In: *16th International Congress of Phonetic Sciences*, Saarbrucken, Germany. 1621-1624.

Nishi, K. & Kewley-Port, D. (2007b). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, 50, 1496-1509.

Nishi, K. & Kewley-Port, D. (2008). Non native speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research*, 51, 1480-1493.

Nittrouer, S. (2001). Challenging the notion of innate phonetic boundaries. *Journal of the Acoustical Society of America*, 110, 1598-605.

Nobre-Oliveira D. (2007). Effects of Perceptual Training on the Learning of English Vowels in Non-native Settings. *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. 382-389.

Nolan, F. (1992). The Descriptive Role of Segments: Evidence from Assimilation. *Physiological and Perceptual Study*. Lund: Lund University Press.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115, 39–57.

O'Grady, W. D. & Dobrovolsky, M. B. (1987). Contemporary linguistic analysis: An introduction. Toronto: Copp Clark Pitman Ltd.

Ortega-Llebaria, M., Faulker, A., & Hazan, V. (2001). Auditory vs. Visual L2 speech perception: effects of visual cues and acoustic-phonetic context for Spanish learners of English. *Proceedings AVSP-2001*. 149-154.

Pattamadilok, C., Knierim, I.N., Kawabata-Duncan, K.J., & Devlin, J.T. (2010). How does Learning to Read Affect Speech Perception? *The Journal of Neuroscience*. 30(25). 8435-8444.

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The orthographic consistency effect in the recognition of French spoken words: an early developmental shift from sublexical to lexical orthographic activation. *Applied Psycholinguistics*, 30, 441– 462.

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, 22, 700 –726.

Paulesu, E. (2006). On the advantage of 'shallow' orthographies: number and grain size of the orthographic units or consistency per se?. *Journal Compilation*. Blackwell Publishing Ltd. 443-444.

Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic influences in spoken word recognition: the consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin Review*, 16, 363–368.

Penfield, W. & Roberts, L. (1959). Speech and brain mechanisms. Princeton, NJ: Princeton University Press.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, 19, 170-182.

Perfetti, C. A. (1984). Reading ability. New York: Oxford University Press.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 139-156.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 572-584.

Pine, J. M. (1994). The language of primary caregivers. In: C. Gallaway & B. J. Richards (Eds.). *Input and Interaction in Language Acquisition*. Cambridge University Press, 15-37.

Pisoni, D. B., Aslin, R. N., Perey, A. J. & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 297-314.

Pisoni, D. B., Lively, S. E. & Logan, J. S. (1994). Perceptual learning of non-native speech contrasts: Implications for theories of speech perception. In J. Goodman & H. C. Nusbaum (Eds.), *Development of speech perception: The transition from speech sounds to spoken words.* Cambridge, MA: MIT Press. 121 – 166.

Pitt, M. A. (1995). The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1037–1052.

Plaut, D. C. & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language*. Mahwah, NJ: Erlbaum. 381-415.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103,* 56-115.

Plaut, D.C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, 12(5/6), 765-805.

Plaut, D.C. (2005). Connectionist approaches to reading. In M.J. Snowling and C. Hulme (Eds.), *The science of reading: A handbook.* Blackwell: Oxford. 24-38.

Polka, L. & Werker, J.F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology and Human Perception and Performance*, 20, 421–435.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America*, 89, 2961-2977.

Polka, L. (1992). Characterizing the influence of native experience on adult speech perception. *Perception and Psychophysics*, 52, 37-52.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America*, 97, 1286-1296.

Polka, L., Colantonio, C. & Sundara, M. (2001). A cross-language comparison of /d/ - /ð/ perception: Evidence for a new developmental pattern. Journal of the Acoustical Society of America, 109, 2190-201.

Potter, M. C., So, K. F., Von Eckhardt, B. & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23, 23-38.

Potter, R.K., Steinberg, J.C. (1950). Toward the specification of speech. *Journal of the Acoustical Society of America*, 22 (6), 807-820.

Price, C. J. & Mechelli, A. (2005). Reading and reading disturbance. *Current Opinion in Neurobiology*, 15(2), 231-238.

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, 197, 335-359.

Pruitt, J. S. (1993). Comments on "Training Japanese listeners to identify english /r/ and /l/: A first report" [J. S. Logan, S. E. Lively, & D. B. Pisoni, J. Acoust. Soc. Am. 89, 874-886 (1991)]. *Journal of the Acoustical Society of America*, 94, 1146-1147.

Pruitt, J. S. (1995). Perceptual training on Hindi dental and retroflex consonants by native English and Japanese speakers. *Journal of the Acoustical Society of America*, 97, 3417.

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, 1684-1696.

Pytlyk, C. (2012). Shared Orthography: Do shared written symbols influence the perception of L2 sounds?. *The Modern Language Journal*. 95 (iv), 541-557.

Rack, J., Snowling, M. & Olson, R. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly,* 27, 29-53.

Ramers, K. H. (1988). *Vokalquantit¨at und -qualit¨at im Deutschen*. Linguistische Arbeiten 213. Tübingen: Niemeyer.

Rastle, K. & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonwords reading. *Psychonomics Bulletin Review*, 5, 277-282.

Rastle, K. & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and non-words, *Journal of Memory and Language, 42*(3), 342-364.

Rauber, A. S., Escudero, P., Bion, R. A. H. & Baptista, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. In: *Proceedings of the Interspeech 2005*, 2913–2916.

Reich. P.A. (1986). Language Development. Englewood Cliffs, N.J: Prentice Hall.

Rivera-Gaxiola. M., Garcia-Sierra, A., Lara-Ayala, L., Cadena, C., Jackson-Maldonado, D., & Kuhl, P. K. (2012). Event-related potentials to an English/Spanish syllabic contrast in Mexican 10-13 month-old infants. *ISRN Neurology*. Vol. 2012, Article ID 702986.

Roach, P. (1991). English Phonetics and Phonology –A practical Course. 2nd Edition. Cambridge University Press.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press. 171-204

Rogers, H. E. (1991). Theoretical and practical phonetics. Mississauga, ON: Copp. Clark Pitman Ltd.

Rosch, E. (1978). Principles of categorization. In: Rosch, E. & B.B. Lloyd (Eds.), *Cognition and Categorization.* Hillsdale, N.J.: Erlbaum.

Saffran, J., Aslin, R., & Newport, E. (1996a). Statistical learning by 8-month old infants. *Science*, 274, 1926–1928.

Saffran, J.R. (2003). Statistical Language Learning: Mechanisms and Constraints. Blackwell Publishing Inc. 12 (4). 110-113.

Saffran, J.R., Newport, E.L. & Aslin, R.N. (1996b*). Word Segmentation: The Role of Distributional Cues. Journal of Memory and Language*, 35, 606-621.

Samaras, M. (1974). Study of the Influence of the Accent on Modern Greek Vowel Formant Structure. *Bulletin de l'Institut de Phonétique de Grenoble,* 3, 129-138.

Schneider,W., Eschman, A. & Zuccolotto, A. (2002a). E-Prime User's Guide. Pittsburgh: Pychology Software Tools Inc.

Schneider,W., Eschman, A. & Zuccolotto, A. (2002b). E-Prime Reference Guide. Pittsburgh: Pychology Software Tools Inc.

Schwartz, A.I., Kroll, J.F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive processes*, 22 (1), 106-129.

Scovel, T. (1988). A time to speak: A Psycholinguistic Inquiry into the Critical Period for Human language. Rowley, MA: Newbury House.

Scovel, T. (2000). A critical review of the Critical Period Hypothesis. *Annual Review of Applied Linguistics,* 20, 213-223.

Scovel. T. (2006). Age, acquisition and accent. In: Age in L2 Acquisition and Teaching. By Christian Abello-Contesse, Ruben Chacon-Beltran. M. Dolores Lopez-Jimenez, M. Mar Torreblanca-Lopez (Eds.). Bern: Peter Lang. 31-48.

Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568.

Seidenberg, M. S. & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology of Human Learning*, 5, 546-554.

Seidl, A. & Johnson, E. K. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, 9(6), 565 – 573.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.

Sfakianaki, A. (2002). The acoustic characteristics of Greek vowels produced by adults and children. Makri-Tsilipakou, M. (Ed.) *Selected Papers on Theoretical and Applied Linguistics*. From the 14th International Symposium on Theoretical and Applied Linguistics. Aristotle University of Thessaloniki. 383-394.

Share, D.L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, 55 (2), 151-218.

Singleton, D. (2001). Age and second language acquisition. *Annual Review of Applied Linguistics*, 21, 77-91.

Sluijter, A.M.C. (1995). Phonetic correlates of stress and accent. Holland Academic Graphics. The Hague.

Smith, F. (2004). Understanding reading (6th ed.). Mahwah, NJ: Lawrence Erlbaum.

Snowling, M. & Frith, U. (1981). The role of sound, shape and orthographic cues in early reading. *British Journal of Psychology*. 72, 83-87.

Stadthagen-Gonzalez, H. & Davis, C.J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. Behavior Research Methods, 38 (4), 598-605.

Stanovich, K. E. (1988). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia*, 38, 154-177.

Steinberg, D. D. & Sciarini, N.V. (2006). An Introduction to Psycholinguistics (2nd Edition). Pearson Education Limited.

Steinberg, D. D. (1993). An introduction to Psycholinguistics. London: Longman.

Stern, D.N., Spieker, S., Barnett, R.K., & MacKain, K. (1983). The prosody of maternal speech: infant age and context related changes. *Journal of Child Language*, 10(1), 1–15.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, 8, 185–190.

Stone, G. O., Vanhoy, M., & van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *The Journal of Memory and Language*, *36*, 337–359.

Strange, W. & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics*, 36, 131–145.

Strange, W. & Jenkins, J. J. (1978). Role of linguistic experience in the perception of speech. In R. D. Walk & H. L. Pick (Eds.), *Perception and experience* (pp. 125-169). New York: Plenum Press.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In: W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research.* Baltimore: York Press. 3-45.

Strange, W. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311–344.

Strange, W. (2010). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*. Available online: http://www.sciencedirect.com/science/article/pii/S0095447010000689.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S.A., Nishi, K. & Jenkins, J.J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics,* 26, 311-344.

Sumby, W.H. & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212–215.

Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R. & Munhall, K.G. (2008). Training English listeners to perceive phonemic length contrasts in Japanese. *Journal of the Acoustical Society of America*, 123, 397-413.

Takagi, N., & Mann, V. A. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics*, 16, 379–405.

Tanenhaus, M. K, Flanigan, H. P., & Seidenberg, M. S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory and Cognition*, 8, 513–520.

Telkemeyer S., Rossi S., Koch S. P., Nierhaus T., Steinbrink J., Poeppel D., Obrig H., & Wartenburger I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *Journal of Neuroscience*, 29, 14726–14733.

Thiessen, E. D. & Saffran, J. R. (2003). When cues collide: Statistical and stress cues in infant word segmentation. *Developmental Psychology*, 39, 706-716.

Thompson, M. & Hazan, V. (2010). The impact of visual cues and lexical knowledge on the perception of a non-native consonant contrast for Colombian adults. *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010*. 493-498.

Tomasello, M. (2003). Constructing a language. Cambridge, MA: Harvard University Press.

Trainor, L.J., & Desjardins, R.N., (2002). Pitch characteristics of infant directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review*, 9 (2), 335–340.

Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology, 33*, 771–780.

Trueswell, J.C., Sekerina, I., Hill, N.M. & Logrip, M.L. (1999a). The kindergartenpath effect: studying on-line sentence processing in young children. *Cognition*, 73, 89-134.

Trueswell, J.C., Sekerina, I., Hill, N.M. & Logrip, M.L. (1999b). The development of on-line language comprehension abilities in children. *Language, Minds, & Brains: Studies in Languages*, 34, pp. 209-215.

Tseva, A. (1989). Structure Acoustique des Voyelles du Grec Moderne: Où en sont les Connaissances?. *Bulletin de l'Institut de Phonétique de Grenoble,* 17(18). 87-107.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: language-specific developmental changes in speech perception abilities. ICSLP-1994, 1695-1698.

Turvey, M.T., Feldman, L.B., and Lukatela, G. (1984). The Serbo-Croatian orthography constraints the reader to a phonologically analytic strategy. In: L. Henderson (Ed.), *Orthographies and Reading: Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Urponen, M. I. (2004). Ultimate attainment in postpuberty second language acquisition. Unpublished doctoral dissertation, Boston University, Boston, MA, U.S.A.

Uther, M., Singh, P., Zipitria, I. & Uther, J. (2005). 'MAC: An adaptive, perception-based speech remediation s/w for mobile devices'. Artificial Intelligence in Education (AIED) workshop on language tutoring, IOS Press.

Uther, M., Uther, J., Athanasopoulos, P., Singh, P. & Akahane-Yamada, R. (2007). Mobile Adaptive CALL (MAC): A lightweight speech-based intervention for mobile language learners. Proceedings of Interspeech 2007 conference, Antwerp, Belgium. 2329-2332.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458-483.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10, 159-164.

Wagner, R. K. & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192-212.

Walley, A. & Flege, J. (1999). Effects of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, 27, 307-332.

Wang, M., Park, Y. J. & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross language and orthography transfer. *Journal of Educational Psychology,* 98, 148–158.

Wang, M., Perfetti, C. A. & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross language and writing system transfer. *Cognition*, 97, 67–88.

Wang, Y., Jongman, A. & Sereno, J.A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America,* 113, 1033-1043.

Wang, Y., Spence, M.M., Jongman, A. & Sereno, J.A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America,* 106, 3649-3658.

Ward, J. (2010). The Student's Guide to Cognitive Neuroscience (2nd Ed). Psychology Press.

Weber-Fox, C., & Neville, H. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In: D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates. 23-38.

Weekes, B.S. (1997). Differential effects of number of letters on word and non-word latency. *Quarterly Journal of Experimental Psychology*, 50A, 439-456.

Wells, J. C. (1962). A Study of the Formants of the Pure Vowels of British English. MA dissertation, UCL. http://www.phon.ucl.ac.uk/home/wells/formants/index.htm.

Werker, J. & Lalonde, C.E. (1988). Cross-Language Speech Perception: Initial Capabilities and Developmental Change. *Developmental Psychology*, 24 (5), 672-683.

Werker, J. F. (1995). Exploring developmental changes in cross-language speech perception. In: D. Osherson (Series Ed.) & L. Gleitman & M. Liberman (Vol Eds.). *An invitation to cognitive science, Part I: Language*. Cambridge, MA: MIT Press. 87-106.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.

White, L. & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12, 238-265.

Wiik, K. (1965). Finnish and English vowels. Annales Universitatis Turkuensis, Series B, 94. Turku, Finland: University of Turku.

Wilson, M.D. (1988) The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, 20(1), 6-11.

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: evidence from children learning to read German. *Journal of Experimental Child Psychology*, 61, 80-90.

Wode, H. (1983). Phonology in L2 acquisition. In H. Wode (Ed.), *Papers on language acquisition, language learning and language teaching.* Heidelberg, Germany: Groos. 175-187.

Woodward, J. C. (1973). Some characteristics of pidgin sign English. *Sign Language Studies*, 3, 39-59.

Wydell, T. N. & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, 70, 273-305.

Wydell, T. N., Vuorinen, T., Helenius, P. & Salmelin, R. (2003). Neural correlates of letter-string length and lexicality during reading in a regular orthography. *Journal of Cognitive Neuroscience*, 15(7), 1052-1062.

Yamada, J. (1991). The discrimination learning of the liquids /r/ and /l/ by Japanese speakers. *Journal of Psycholinguistic Research*, 20 (1), 31-46.

Yamada, R. A. (1993). Effects of extended training on /r/ and /l/ identification by native speakers of Japanese. *Journal of the Acoustical Society of America*, 93, 2391(A).

Ylinen, S., Shestakova, A., Alku, P. & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and non-speakers of Finnish, *Language and Speech*, 48, 313–338.

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.

Yoshida, K. A., Pons, F., Cady, J. C. & Werker, J. F. (2006). Distributional learning and attention in phonological development. Paper presented at Int. Conf. on Infant Studies, Kyoto Japan.

Yu, C., Ballard, D. H. & Aslin, R. N. (2005). The role of embodied intention in early lexical acquisition. Cognitive Science, 29, 961–1005.

Zhang, Y., Kuhl, P.K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage,* 26, 703-720.

Ziegler, J. & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, 131 (1), 3-29.

Ziegler, J. & Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science,* 9 (5), 429-453.

Ziegler, J. C., Ferrand, L. (1998). Orthography shapes the perception of speech: the consistency effect in auditory word recognition. *Psychonomic Bulletin Review,* 5, 683– 689.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science,* 12(5), 379-384.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: where do we stand after more than a decade? Journal of Experimental Psychology: Learning Memory & Cognition, 34, 643– 661.

Ziegler, J. C., Stone, G. O. & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behaviour Research Methods, Instruments, & Computers*, 29, 600–618.Oller, J.W., & Ziahosseiny, S.M. (1970). The contrastive analysis hypothesis and spelling errors. *Language Learning* , *20*, 183–189.