UNIVERSITAT DE BARCELONA

The Role of Cognitive Ability in the Acquisition of Second Language Perceptual Phonological Competence

Elena Safronova

THE ROLE OF COGNITIVE ABILITY IN THE ACQUISITION OF SECOND LANGUAGE PERCEPTUAL PHONOLOGICAL COMPETENCE

BY

ELENA SAFRONOVA

DISSERTATION

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Department of English Studies and

Modern Languages and Literatures

Doctoral Program in Linguistic,

Literary and Cultural Studies

University of Barcelona

Director: Dr. Joan C. Mora

2016

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

Elena Safronova

Barcelona, September 2016

Dedication

This dissertation is lovingly dedicated to my parents, Olga Safronova and Leonid Safronov. Their advice, support, and constant love have sustained me throughout my life and made this dissertation come true.

Acknowledgments

This dissertation would not have seen the light of day without help of many people. I would like to express my deepest gratitude to my supervisor, Dr. Joan C. Mora. He has inspired my interest in and given me immense knowledge about L2 psycholinguistics and phonology. I would have never made it without his support and invaluable guidance. I would also like to thank the members of my follow-up committee Dr. Natalia Fullana, Dr. Teresa Navés and Dr. Llorenç Comajoan Colomé for their insightful comments which helped considerably to improve this dissertation.

I am very grateful to Dr. Carmen Muñoz for giving me an opportunity to work as a research assistant at the Department of English and German Philology at the University of Barcelona, learn from her and gain invaluable experience in conducting high-quality research. My sincere thanks goes to all teachers in the Master's and PhD programs in Applied Linguistics for sharing their knowledge and introducing me to the fascinating world of SLA research. I would like to thank students of the University of Barcelona who took part in my study.

I would like to acknowledge the funding (BE-DGR 2012 BE00359) provided by the Catalan Ministry of Economy and Knowledge and the Agency for the Management of University and Research Grants (AGAUR), which has supported my research stay at Indiana University, from which this PhD work has greatly benefited.

I would like to express my sincere appreciation to Dr. Isabelle Darcy for her care, important suggestions and help with the data collection during my research stay at her L2 psycholinguistic lab at Indiana University. She gave me an opportunity to meet with the members of her research team who gave me their priceless feedback which helped make significant progress with the dissertation. Special thanks goes to Dr. Jeff Holliday for his advice and help with data collection instruments. I am very grateful to Dr. Darcy's PhD and undergraduate students who participated in my data collection.

I would like to thank my colleagues and friends I was lucky to meet during my studies at the University of Barcelona. They are Mayya Levkina, Elena Markina, Eva Cerviño-Povedano, Verónica Rodríguez and Hanna Kivistö-de Souza. We discussed our research, exchanged ideas and helped each other. These people made my time in Barcelona so memorable. I am so very grateful to Eva for sharing her phonological memory test and our encouraging conversations. Special thanks goes to Mayya Levkina for the great time we had working and spending spare time together, for her support and for being my friend.

My deepest thanks goes to my parents and my sister who have given me infinite love, courage and guidance throughout my life. I would have never been who I am without them. I wish to thank with all my heart my husband Viktor Kravchenko for broadening my horizons, encouraging and supporting me (both morally and intellectually) throughout this work. Finally, I wish to thank my son Daniel for being the strongest motivation to finish this dissertation.

Abstract

Many second language (L2) learners find L2 pronunciation difficult and experience perception and production problems leading to accented speech. There exists a great inter-learner variation in L2 phonological acquisition even among learners who have been exposed to the L2 since childhood. Such individual differences have been the focus of much second language acquisition (SLA) research. Relatively little attention has been paid to the role of learners' cognitive ability in the acquisition of L2 perceptual phonological competence. This dissertation seeks to fill this gap by addressing the following question: To what extent does cognitive ability contribute to learners' acquisition of L2 sounds? We hypothesized that phonological short-term memory, acoustic short-term memory and attention control facilitate L2 learners' acquisition of perceptual phonological competence.

A group of 45 adult Catalan-Spanish bilingual learners of English were asked to participate in a battery of L2 perception tests measuring their L2 vowel perception and cognitive tests assessing their attention control ability and short-term memory capacity for speech sounds. The target L2 sounds were those comprised in the English vowel contrasts /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/. The learners' ability to perceive a cross-language phonetic distance between these L2 vowels and those in their native language (L1) and their success in establishing new phonetic categories for L2 vowels were assessed by means of a perceptual assimilation task and a categorical vowel discrimination task, respectively. The contribution of phonological short-term memory, acoustic short-term memory and attention control was examined by relating the outcome measures of cognitive tests to those of learners' L2 vowel perception.

Overall, the results obtained partly confirmed the hypothesis. Short-term memory capacity for phonological and acoustic information and attention control ability significantly contributed to explaining the variance in learners' perception of L2 sounds. Attention control and acoustic short-term memory were related to learners' perception of cross-language phonetic distance. Contrary to our predictions, larger phonological short-term memory and acoustic short-term memory capacities were associated with lower degree of perceived phonetic distance between L2 and L1 sounds. In addition, lower attention control was related to faster and more accurate discrimination of L2 sounds. Taken together, our findings indicate that cognitive ability plays a role in L2 learners' acquisition of perceptual phonological competence.

Keywords: individual differences, L2 perceptual phonological competence, phonological short-term memory, acoustic short-term memory, attention control.

Resumen

La mayoría de los aprendices de segundas lenguas (L2) tienen dificultades en la pronunciación de una L2 y tienen problemas con la percepción y la producción del habla que resultan en un acento extranjero. Existe una gran variación entre aprendices en la adquisición fonológica de la L2, incluso entre los que han estado expuestos a su L2 desde la infancia. Tales diferencias individuales han sido el foco de investigación en la adquisición de segundas lenguas. Se ha prestado relativamente poca atención al papel de la capacidad cognitiva de los aprendices en la adquisición de la competencia perceptiva en una L2. Esta tesis doctoral pretende llenar este vacío, abordando la siguiente pregunta: ¿En qué medida la capacidad cognitiva contribuye a la adquisición de los sonidos de una L2? Nuestra hipótesis es que la memoria fonológica a corto plazo, la memoria acústica a corto plazo y el control de la atención facilitan la adquisición de la competencia perceptiva de los aprendices de una L2.

Un grupo de 45 aprendices adultos catalán-español de inglés participaron en una batería de pruebas de percepción del habla midiendo su percepción vocalica en la L2 y de pruebas cognitivas midiendo su control de la atención y la capacidad de memoria a corto plazo para los sonidos del habla. Las categorías fonéticas de la L2 en el foco eran las comprendidas por los contrastes entre vocales inglesas /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ y /ʌ/-/æ/. La capacidad de los aprendices de percibir una distancia fonética entre estas vocales y las de su lengua materna (L1) y su éxito en establecer vocales de la L2 se evaluaron por medio de una tarea de asimilación perceptiva y una tarea de discriminación de vocales, respectivamente. La contribución de la memoria fonológica, la memoria acústica y el control de la atención se examinó relacionando las medidas cognitivas con la percepción vocalica en la L2.

En general, los resultados obtenidos en parte confirman la hipótesis. La capacidad de memoria a corto plazo para la información fonológica y acústica y el control de la atención contribuyeron significativamente a explicar la variación en la percepción de los sonidos de una L2. El control de la atención y la memoria acústica resultaron ser los mejores predictores del grado de la distancia fonética percibida entre los sonidos de la L2 y la L1. Contrariamente a nuestras predicciones, la memoria fonológica y la memoria acústica se asociaron con un mayor grado de similitud percibida entre los sonidos de la L2 y la L1. Además, una menor capacidad del control de la atención se relacionó con una percepción más precisa de oposiciones vocalicas de la L2. En conjunto, los resultados sugieren que la capacidad cognitiva desempeña un papel en la adquisición de la competencia perceptiva en una L2.

Palabras clave: las diferencias individuales, la competencia perceptiva en la L2, la memoria fonológica, la memoria acústica, el control de la atención.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

All our knowledge is the offspring of our perceptions.

—LEONARDO DA VINCI (1452-1519)

1.1 Motivation of the dissertation

One of the goals people learning a second/foreign language (L2) pursue is to understand L2 speech and to make themselves understood by an L2 speaker[1]. The ability to accurately perceive and pronounce L2 sounds constitutes one of the L2 phonological aspects learners have to master in order to achieve this goal. It has long been well established that many learners struggle to accurately perceive and produce L2 sound contrasts that do not signal differences in word meaning (henceforth *phonological contrasts*) in their native language (L1; Flege & Hillenbrand, 1987; Goto, 1971; Mack, 1989; Pallier, Bosch, & Sebastián-Gallés, 1997). Yet, there exist learners who accomplish this task with greater ease than others and succeed in attaining a high level of L2 phonological competence. Specifically, some learners are better able to accurately perceive and produce L2 sounds than others. The explanation of this phenomenon requires joint efforts of the research in second language acquisition (SLA) and psycholinguistics. This dissertation addresses the

[1] In the present dissertation the term "second language" will be used to refer to a language learned in a naturalistic or formal instructional setting after an individual's mother tongue(s) has been acquired.

issue of inter-learner variation in L2 phonological competence by examining the role of L2 learners' cognitive skills in the acquisition of L2 phonological contrasts[2].

According to the *Common European Framework of Reference for Languages: learning, teaching, assessment* (2002: 116-117) one of the core components of the phonological competence is "the knowledge of, and skill in perception and production" of the speech sounds (i.e. vowels and consonants) and the phonetic features distinguishing them. The present dissertation is primarily concerned with L2 perceptual phonological competence, by which we understand the extent to which an L2 learner has established accurate long-term representations for L2 sounds used contrastively in the L2 to convey differences in meaning (i.e. L2 phonetic categories).

There are two reasons that explain the focus of this dissertation on perceptual phonological competence. One of them concerns the role accurate perception of speech sounds plays in successful communication. L2 learners' inability to correctly identify the speech sounds of which words are composed may cause miscomprehension of a spoken message (McAllister, 1997; Strange & Shafer, 2008). For example, if an L2 learner does not accurately perceive the English /i/-/ɪ/ phonological contrast, which distinguishes words like *sheep* /ʃiːp/ and *ship* /ʃɪp/, (s)he may misinterpret an interlocutor's message like *I took a picture of a black ship yesterday*. In other words, the L2 learner may incorrectly understand that the object in the picture is an ovine rather than a marine vessel. The other reason is that the ability to accurately perceive L2 sounds has been found to facilitate their production, leading to lower degrees of foreign accent (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Flege, 1995; Hazan, Sennema, Iba, & Faulkner, 2005).

Therefore, achieving a high level of perceptual phonological competence may be considered one of the important tasks for L2 learners. The fact that we still do not fully understand the sources of L2 learners' difficulties in accomplishing this task indicates

[2]The distinction between the terms "L2 learning" and "L2 acquisition" (Krashen, 1981) is not discussed in this dissertation. The term "learning" will be used to emphasize the process of developing L2 skills, whereas the term "acquisition" will be understood as a successful mastery of those L2 skills (Flege, 1988).

the need for further research on L2 speech perception. This will help better understand the mechanisms underlying successful L2 speech learning[3]. Moreover, the research outcomes can be used to design effective training programs aimed at improving L2 learners' pronunciation skills.

1.2 Individual differences in L2 speech learning

It has been widely observed that some learners manage to attain a native-like level of L2 phonological competence whereas others do not. A substantial body of SLA studies has been conducted in order to identify factors contributing to successful L2 speech learning (see Piske, MacKay, & Flege, 2001; Purcell & Suter, 1980, for a review). The following factors have been identified in previous research as making a significant contribution: the age of onset of L2 learning (Baker, Trofimovich, Flege, Mack, & Halter, 2008; Flege & MacKay, 2004; Flege, Yeni-Komshian, & Liu, 1999; Flege, MacKay, & Meador, 1999; Long, 1990), the L2 learners' L1 background (Flege, Bohn, & Jang, 1997), the length of residence in an L2 environment (Flege et al., 1997; Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Purcell & Suter, 1980), the amount of L1 and L2 use (Flege & MacKay, 2004), the quantity and quality of L2 input (Flege, 2009; Flege & Liu, 2001; Moyer, 2009) and L2 lexical knowledge (Bungaard-Nielsen, Best, & Tyler, 2011). Despite a growing number of studies supporting the role of the above mentioned factors in L2 speech learning, the inter-learner variation in L2 phonological competence remains even when those factors are controlled.

1.2.1 Age, experience and proficiency

Research on the age factor has shown that "the earlier the better" in second language acquisition (Flege, 1987; Long, 1990). However, despite age being considered a strong con-

[3]The term "L2 speech learning" will be referred to as the process of gaining knowledge about the L2 phonological system through formal and/or naturalistic exposure.

tributor, there exists a large inter-learner variation in L2 phonological acquisition among late as well as early learners. On the one hand, there are exceptionally talented adult L2 learners who have manged to acquire native-like L2 pronunciation skills (Birdsong, 2007; Bongaerts, 1999; Ioup, Boustagui, El Tigi, & Moselle, 1994; Moyer, 2014). L2 learners who started learning an L2 in childhood have been found not only to have a detectable foreign accent, but also to differ from native speakers in L2 speech perception (Flege, Yeni-Komshian, & Liu, 1999; Flege et al., 2006; Højen & Flege, 2006). Moreover, studies on L2 speech learning in a formal instructional setting demonstrate that older L2 learners may be better able to accurately perceive L2 phonological contrasts and have a lower degree of foreign accent than younger learners (Fullana, 2006; García Lecumberri & Gallardo, 2003).

There is also evidence that L2 experience, operationalized as length of residence in an L2-speaking country, may not be a robust predictor of accurate L2 speech perception and production (McAllister, 2001; Moyer, 1999). Several studies with bilingual populations have shown that the extensive exposure to high-quality L2 input since early childhood does not guarantee target-like perception and production of L2 vowel contrasts (Pallier et al., 1997; Sebastián-Gallés, Echeverría, & Bosch, 2005; Sebastián-Gallés & Soto-Faraco, 1999).

As regards L2 proficiency, often being measured in terms of L2 learners' vocabulary knowledge, previous SLA studies have provided contradictory results as to its role in L2 speech learning. Bungaard-Nielsen et al. (2011) found that L2 learners with larger vocabularies perceive many L2 vowel contrasts more accurately than those with smaller vocabularies. However, the study showed that learners' differences in vocabulary size may not be related to the perception of L2 sounds which are perceptually too easy or too difficult to discriminate. A number of recent studies found that L2 proficiency was unrelated to inter-learner variation in accurate perception and production of L2 phonological contrasts (Cerviño-Povedano & Mora, 2015; Darcy, Mora, & Daidone, 2014; Mora & Safronova, submitted). Taken together, previous research findings suggest that inter-learner variation in L2 speech learning may exist irrespective of such important factors as age of onset of L2 learning, experience and proficiency.

1.2.2 L1 influence on L2 speech learning

The influence of the previously acquired L1 has long been observed in many aspects of learners' L2 such as the lexicon, syntax and phonology (see Major, 2001; Odlin, 1989, for a review). In L2 phonology, L1 influence is evident in L2 speakers' foreign accent, which often allows a native listener to identify the non-native speaker's L1 background. In research on L2 speech speech perception, L1 influence has been studied in relation to the effects of previous experience with L1 speech on the accurate perception of L2 sounds (Best & Tyler, 2007; Flege, 1995; Kuhl & Iverson, 1995). The presence of L1 influence in L2 speech perception is the consequence of the phenomenon of categorical perception.

The origin of this phenomenon lies in humans' innate capacity to classify items in the world into categories. Perceived similarities and differences among items allow for grouping some items into the same category while placing others into a different category (Harnad, 2003; see also Lakoff, 1987; Rosch, 1978). The way categorical perception influences learning of L2 sounds can be best captured by Trubetzkoy's (1969) "phonological sieve" metaphor. According to Trubetzkoy, one's earlier acquired L1 phonological system functions as a perceptual "sieve" through which L2 sounds are filtered and become categorized on the basis of perceived similarity to L1 categories.

There are several L2 speech learning models which are concerned with the prediction and explanation of difficulties in learning L2 sounds (see Boersma & Hamann, 2009; Bohn, 2002, for a review of L2 speech perception models). Among them are the Perceptual Assimilation Model (PAM-L2: Best & Tyler, 2007), the Speech Learning Model (SLM: Flege, 1995, 2007), the Native Language Magnet model (NLM: Kuhl & Iverson, 1995; Kuhl et al., 2008) and the Second Language Linguistic Perception model (L2LP: Escudero, 2009; van Leussen & Escudero, 2015). A common tenet of these models is that the previously acquired L1 phonological system influences learners' perception of L2 sounds, which governs the development of L2 phonological system.

These models contend that L2 learners' perception of cross-language phonetic similarity, that is, a degree of perceived distance between L1 and L2 sounds, determines whether L2 sounds become assigned to newly created L2 phonetic categories or equated with already existing L1 categories. A phonetic category is defined as a long-term memory representation for a wide range of different speech sounds identified as being the same, "despite auditorily detectable differences between them along dimensions that are not phonetically relevant" (Flege, 1995: 244). If L2 learners fail to perceive phonetic differences between an L2 and an L1 sounds, the L2 sound will be perceptually mapped onto the L1 category. For instance, for Catalan learners of English, English /ʌ/ will be perceptually mapped onto Catalan /a/. In this case a new phonetic category for the L2 sound (English /ʌ/) would not be established (Flege, 1995). Moreover, if two or more L2 sounds are perceived similarly to the same L1 sound (for instance, English /ɑ/-/ʌ/ both perceived as Catalan /a/), both L2 sounds will be assigned to the same L1 category. In this scenario, L2 learners will have difficulties in learning such L2 sound contrasts (Best & Tyler, 2007; van Leussen & Escudero, 2015), which might further lead to miscomprehension of L2 speech. In other words, the establishment of distinct phonetic categories for L2 sounds like English /i/-/ɪ/ by a learner whose L1 lacks this phonological contrast is crucial for his/her ability to distinguish differences in word meaning, such as the difference in meaning between *sheep* /ʃiːp/ and *ship* /ʃɪp/.

L2 learners can establish L2 phonetic categories, but this process is complex and requires an ability to discern at least some phonetic differences between L2 and L1 sounds as well as between contrasting L2 sounds (Best & Tyler, 2007; Flege, 1995). In other words, in order to acquire L2 sounds, learners must be able to perceive a cross-language phonetic distance between the sounds in their L1 and L2. The L2 speech learning models, which will be discussed in detail in Chapter 2, do not suggest what mechanisms and factors may contribute to this ability.

This dissertation focuses on learners' individual differences in the ability to establish new phonetic categories for L2 sounds. We assumed that this ability is determined by L2 learners' perception of phonetic distance between L2 and L1 sounds, and indexed by the accurate perception of L2 phonological contrasts. As discussed earlier in the chapter, factors such as age of L2 learning, L2 experience and L2 proficiency were not always found to explain all of the inter-learner variation in L2 speech learning. Therefore, it is important to investigate L2 learners' differences in perceptual phonological competence from the perspective of their individual differences in cognitive ability.

1.2.3 Cognitive ability in L2 speech learning

In order to account for inter-learner differences in L2 acquisition, SLA research has focused on a psycholinguistic approach emphasizing the role of information processing mechanisms and cognitive skills in L2 learning (Dörnyei, 2005; Dörnyei & Skehan, 2003; Miyake & Friedman, 1998; Skehan, 1998; Robinson, 2002, 2012; Skehan, 2002, 2012). The psycholinguistic approach to SLA views L2 learning as a cognitively demanding process that requires L2 learners to make use of their cognitive abilities, such as memory and attention. Inter-learner variation in rate and success of L2 acquisition has been attributed to L2 learners' differences in L2 aptitude, that is, variation in how efficient individuals are in making use of the cognitive resources the task of learning an L2 requires (Robinson, 2005).

As regards cognitive abilities underlying L2 speech learning, working memory (Darcy, Park, & Yang, 2011, 2015), phonological short-term memory (Aliaga-Garcia, Mora, & Cerviño-Povedano, 2011; Cerviño-Povedano & Mora, 2011, 2015; MacKay, Meador, & Flege, 2001), acoustic short-term memory (Mora & Safronova, submitted; Safronova & Mora, 2012a; Tanaka & Nakamura, 2004), attention (Darcy et al., 2014; Francis, Baldwin, & Nusbaum, 2000; Francis & Nusbaum, 2002; Guion & Pederson, 2007) and inhibition (Darcy, Mora, & Daidone, 2016; Lev Ari & Peperkamp, 2014) have been found to be related to learners' L2 speech perception and production.

Being a mental "workbench" at which necessary information is temporarily stored and processed, working memory is implicated in the execution of complex cognitive tasks such as learning, comprehension and problem solving (Baddeley, 1996a, 2012; Cowan, 2008). In SLA research working memory has been proposed to constitute a central component of L2 learning aptitude (Dogil & Reiterer, 2009; Dörnyei, 2005; Dörnyei & Skehan, 2003; Miyake & Friedman, 1998; Sawyer & Ranta, 2001; Skehan, 2002). A greater working memory capacity is believed to play an important role in the acquisition of L2 speech because it may "allow learners more time to process and learn from the input by maintaining longer access to it, and better storage quality might promote more accurate perception and learning" (Darcy et al., 2015: 63-64).

In particular, SLA research is concerned with two working memory components, namely, the phonological loop and the central executive. The former has been often associated with phonological short-term memory, referred to as a temporary storage of verbal information by the means of subvocal rehearsal. For example, when we have to memorize a car number we need to repeat it using our inner speech. This ability has been found to be related to successful acquisition of L1 and L2 vocabulary (Atkins & Baddeley, 1998; Baddeley, Gathercole, & Papagno, 1998; Papagno & Vallar, 1995; Silbert et al., 2015, among others) and L2 grammar (O'Brien, Segalowitz, Collentine, & Freed, 2006). Regarding L2 phonology, previous studies have provided evidence that phonological short-term memory promotes L2 learners' gains in perceptual acquisition of L2 sounds under conditions of naturalistic exposure (MacKay et al., 2001) and phonetic training (Aliaga-Garcia et al., 2011). The results of these studies suggest that greater phonological short-term memory capacity is related to L2 learners' accurate perception of L2 vowels and consonants and can be a predictor of successful acquisition of L2 sounds.

Learning L2 speech also requires storing and processing acoustic information contained in speech input. In other words, if we have to decide whether two words are produced by the same speaker we need to be able to temporarily store the acoustic information characterizing the speaker's speech sounds (e.g., pitch and intensity). Previous research has suggested that, despite operating primarily at the categorical (i.e. phonological) level, working memory is also capable of processing acoustic information, which makes acoustic memory its potential component (Baddeley, 2007, 2012; Friedrich, 1990; Joseph et al., 2015; Williamson, Baddeley, & Hitch, 2010). Acoustic short-term memory has received considerable attention regarding its role in the discrimination of vowels and consonants on the basis of acoustic information it can temporarily store (Fujisaki & Kawashima, 1970; Pisoni, 1973, 1975). Yet, little is known about the contribution of acoustic short-term memory to L2 speech learning. Only a few studies have investigated the relationship between acoustic short-term memory and L2 speech learning (Mora & Safronova, submitted; Safronova & Mora, 2012a; Tanaka & Nakamura, 2004). Overall, their results indicate that a larger acoustic short-term memory capacity is associated with a more native-like perception and production of L2 sounds.

The central executive component of working memory has been studied in SLA in terms of the role its attention control functions play in language learning. Attention control functions such as selective attention (Francis & Nusbaum, 2002), directing of attention (Guion & Pederson, 2007), inhibition (Darcy et al., 2016; Lev Ari & Peperkamp, 2014) and attention shifting (Darcy et al., 2014; Mora & Safronova, submitted; Safronova & Mora, 2013) have been found to be related to learning of L2 sounds. This dissertation is primarily concerned with attention shifting, that is, the ability to quickly switch the focus of attention back and forth between multiple tasks while maintaining accurate performance in those tasks (Isaacs & Trofimovich, 2011; Monsell, 2003; Segalowitz & Frenkiel-Fishman, 2005). For example, while reading a magazine, we may have to shift our attention once we hear a phone call. An efficient attention-shifting ability will be needed to understand the calling person's message and then get back to the magazine.

This ability may potentially play an important role in L2 speech learning by contributing to L2 users remaining sensitive to differences between L2 sounds while comprehending the incoming speech (Segalowitz, 1997). According to Segalowitz (1997), individual differences in attention control may lead to inter-individual variation in L2 attainment.

Indeed, attention shifting has been found to significantly predict L2 learners' proficiency (Segalowitz & Frenkiel-Fishman, 2005) and to be related to L2 learners' differences in L2 speech perception and production (Darcy et al., 2014; Mora & Safronova, submitted; Safronova & Mora, 2013).

In sum, L2 psycholinguistic studies provide evidence that there is a relationship between cognitive ability and L2 speech learning. The ability to establish new phonetic categories for L2 sounds may require L2 learners' cognitive abilities such as acoustic short-term memory, phonological short-term memory and attention control. Specifically, larger short-term memory capacity for acoustic and phonological information, and efficient ability to shift attention between linguistic dimensions may enhance L2 learners' ability to perceive the acoustic-phonetic distance between L2 and L1 sounds, and between contrasting L2 sounds while processing L2 speech. Still, it is not known to what extent these cognitive skills contribute to learners' perception of cross-language phonetic differences and their acquisition of L2 perceptual phonological competence.

1.3 The present dissertation: aims, research questions and hypotheses

The purpose of this dissertation is to investigate the role of phonological short-term memory, acoustic short-term memory and attention control in the acquisition of L2 perceptual phonological competence. More specifically, this dissertation aims at finding out if L2 learners' ability to perceive a cross-language phonetic distance between L2 and L1 sounds, and their acquisition of contrasting L2 sounds is promoted by such cognitive abilities. The following research questions are addressed:

1. To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' degree of perceived phonetic distance between L2 and L1 vowels?

2. To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' perception of L2 vowel contrasts?

We hypothesize that there exists a relationship between phonological short-term memory, acoustic short-term memory, attention control and L2 learners' acquisition of perceptual phonological competence. Specifically, these cognitive abilities contribute to L2 learners' ability to perceive a cross-language phonetic distance between L2 and L1 vowels, and to their perception of L2 vowel contrasts.

In order to answer the research questions, we assessed L2 learners' vowel perception by means of two L2 vowel perception tasks. Their phonological short-term memory, acoustic short-term memory and attention control were measured through a battery of speech-based cognitive tasks. The learners' linguistic background information was recorded and their vocabulary knowledge was assessed to take into account L2 learning-related factors such as age of onset of L2 learning, amount of L2 exposure and L2 proficiency when analyzing individual differences in the acquisition of L2 perceptual phonological competence.

1.4 Outline of the dissertation

This chapter has introduced the main issues concerning learners' individual differences in the acquisition of L2 phonological competence, which are discussed in detail in the remaining six chapters of this dissertation. The following two chapters provide an overview of previous research outcomes as regards learning of L2 sounds and cognitive skills. Specifically, Chapter 2 discusses the origin of L2 learners' difficulties in the acquisition of L2 perceptual phonological competence by describing the mechanisms involved in speech perception, such as categorical perception and cross-language phonetic similarity. The chapter also reviews current L2 speech learning models.

Chapter 3 presents a review of previous research on individual differences in cognitive ability and L2 speech learning. In particular, the focus is on working memory as the central component of L2 learning aptitude. Baddeley's (2002b, 2012) model of working memory is discussed. Previous research findings on the role of phonological short-term memory and attention control in L2 speech learning are presented. The nature of acoustic short-term memory as a potential component of working memory and its role in the acquisition of L2 perceptual phonological competence are discussed.

Chapter 4 presents the methodology of the study we conducted to estimate the contribution of L2 learners' phonological short-term memory, acoustic short-term memory and attention control to learners' perception of cross-language phonetic distances and their acquisition of L2 perceptual phonological competence. The chapter describes L2 perception and cognitive tasks used to assess the participants' L2 vowel perception, phonological short-term memory, acoustic short-term memory and attention control abilities.

The results obtained are presented in Chapter 5. Chapter 6 evaluates and discusses the results in the light of previous research findings, suggests possible implications and presents the limitations of the study. Chapter 7 concludes this dissertation by outlining its contribution to the field of SLA and proposing directions for future research.

Chapter 2

Issues in L2 speech perception

2.1 Introduction

The acquisition of L2 phonological contrasts is partly determined by learners' success in identifying and discriminating pairs of contrasting L2 sounds. L2 speech perception is affected by L2 learners' previous linguistic experience with their L1 (Best & Tyler, 2007; Flege, 1995). More than 70 years ago one's L1 was compared to a perceptual "sieve" filtering out the acoustic-phonetic properties of L2 sounds (Trubetzkoy, 1969: 51). Since then much L2 speech research has investigated the mechanisms involved in L2 speech perception.

The present chapter aims at reviewing previous research outcomes concerning the aspects of speech perception associated with learners' difficulties in acquiring L2 perceptual phonological competence. This will set the context for a research-based theoretical framework that will be useful in providing a better understanding of the mechanisms underlying L2 speech learning.

2.2 The nature of speech perception

In order to understand what causes L2 learners' difficulties in L2 speech learning it is important to look into the mechanisms involved in L1 and L2 speech perception. This section presents an overview of the modes and levels of human speech perception. This will help explain how L2 speech perception works and how it determines the acquisition of L2 phonological competence.

2.2.1 Modes of speech perception

One of the central distinctions in speech perception theory is that between two main modes of speech perception, namely, the continuous and the categorical modes (Pisoni & Lazarus, 1974; Strange, 2002; Wode, 1994). These two modes are given at birth and "constitute the original innate sensitivities for speech perception" (Wode, 1994: 334). The interaction of the continuous and the categorical modes of perception constitutes a mechanism for L1 and L2 phonological acquisition that remains intact throughout the lifespan.

The continuous mode of speech perception is responsible for encoding acoustic properties of speech sounds. It allows discriminating sounds according to a gradual scale based on acoustic dimensions such as loudness, pitch and voicing. Due to its high sensitivity to low-level acoustic details the continuous mode plays an important role in the perception of acoustic properties of L1 and L2 sounds and in shaping the L1 and L2 phonological systems. Once a phonological system has been established, the perceptual system starts to operate primarily in the categorical mode.

The categorical mode of perception emerges from a fundamental human cognitive capacity to classify what we see, hear and feel into distinct categories on the basis of similarity between the items assigned to a particular category (Lakoff, 1987; Repp, 1984). Categorical perception is considered crucial for the efficient working of an organism as it allows processing of "maximum information with the least cognitive effort" (Rosch, 1978: 28). Categorical perception accomplishes this by minimizing the infinite differences among the stimuli within the same category and emphasizing between-category differences. Therefore, the stimuli within the same category are perceived as more similar than stimuli belonging to different categories.

In speech perception research categorical perception is defined as "the tendency for adult listeners of a particular language to classify the sounds used in their language as one phoneme or another, showing no sensitivity to intermediate sounds" (Kuhl, 2004: 833). Due to its absolute, all-or-none nature, the categorical mode of perception allows speech processing to work rapidly and efficiently (Wode, 1994). The effects of categorization on the perception of speech segments (i.e. vowels and consonants) have long been observed and is well documented by previous studies on speech perception (Beddor & Strange, 1982; Kuhl, 1991; Liberman, Harris, Hoffman, & Griffith, 1957; Miyawaki et al., 1975; Pisoni, 1973, 1975).

For instance, Liberman et al. (1957) found that listeners categorize a continuum of synthetic stop consonants (/b/-/d/-/g/) into sharply distinct categories. The participants in their study were more accurate at discriminating between sounds they had previously identified as belonging to different phonemes (e.g., /b/ vs. /g/) than between sounds that were different realizations of the same phoneme (e.g., two tokens of /b/). There is evidence that the ability to categorize speech sounds into categories emerges very early in life. Young infants have been found to be able to distinguish vowels like /i/-/a/ and /i/-/u/ which constitute two different categories (Trehub, 1973), and pay less attention to the differences within the same category, for instance, between two tokens of /b/ sound (Eimas, Siqueland, Jusczyk, & Vigorito, 1971).

To sum up, in the course of phonological acquisition one's perceptual system establishes categories for speech sounds on the basis of the continuous mode. In order to facilitate the rapid and efficient processing of information the perceptual system starts to operate primarily in the categorical mode ignoring acoustic differences between speech sounds assigned to the same phonetic category. The emergence of the categorical mode, brought about by the development of the L1 phonological system, causes difficulties for subsequent L2 speech learning. While developing the L2 phonological system, learners may ignore acoustic properties of L2 sounds and form L2 phonetic categories on the basis of already established L1 categories. Learning L2 sounds thus requires learners to use both modes when perceiving L2 speech, which involves temporarily storing continuous and categorical information about L2 speech input in memory (the two short-term memory stores are discussed in Chapter 3). An overview of the levels of speech perception may help shed more light on how the L1 phonological system affects L2 phonological acquisition.

2.2.2 Modeling speech perception and spoken word recognition

The process of perceiving and recognizing an incoming spoken word has been modeled as a process consisting of four levels (see Figure1): acoustic, phonetic, phonological (or phonemic) and lexical (Pisoni & Luce, 1987; Ramus et al., 2010; van Leussen & Escudero, 2015). The acoustic level is the level at which the perceptual system conducts a preliminary analysis of the raw acoustic signal represented in terms of acoustic cues, such as spectral structure, frequency, intensity, and temporal attributes. At this level the perceptual system can detect speaker-specific acoustic differences between two tokens of the same speech sound (e.g., [ɑː] in *cot*) produced by two different speakers of American English.

The phonetic level is the level at which the acoustic cues are used to identify speech sounds (or phones) specified by the phonetic properties that distinguish their pronunciation from that of other speech sounds (Atkinson-King, 1980; Pisoni & Luce, 1987). For example, at this level the perceptual system makes use of spectral cues such as height and backness to identify English [ɑː] in *cot* and [ʌ] in *cut* as being different speech sounds This is the first level where the categorization of speech sounds on the basis of their physical differences takes place.

At the phonological level phones are linked to their abstract phonological representations that contrast word meaning in language (i.e. phonemes). In the example given above, English phones [ɑː] and [ʌ] will be mapped onto two distinct English phonemes /ɑ/ and /ʌ/, respectively. In other words, at this level speech sounds are categorized as instances of the respective phonemes which are further connected to the word meanings (*cot* vs. *cut*) at the lexical level. This last level of speech perception allows the listener to comprehend a spoken word.

In the course of L1 phonological acquisition speech perception becomes attuned to the properties of the native language at the phonetic and phonological levels (Ramus et al., 2010). This facilitates accurate, ballistic and effortless L1 speech processing while affecting the perception of L2 sounds. For instance, Catalan does not have a phonological distinction between /ɑ/ and /ʌ/ (Carbonell & Llisterri, 1999). Therefore, the attunement of speech perception to the L1-specific phonetic properties of Catalan may cause Catalan listeners to identify English [ɑː] and [ʌ] as phonetic variants (i.e. allophones) of a single native /a/ vowel (Rallo Fabra & Romero, 2012). This may prevent these English phones from being mapped onto distinct English /ɑ/ and /ʌ/ vowel categories and, as a result, lead to the miscomprehension of a spoken word.

To summarize, successful word recognition requires the establishment of accurate language-specific representations of speech sounds (i.e. phonetic categories). In order for the L1 and L2 phonological systems to be acquired, the continuous and the categorical modes of speech perception must interact so that the language-specific acoustic-phonetic properties of speech sounds can be extracted and used to shape the L1- and L2-specific categories. It appears that the already established L1 phonological system affects the perception of L2 speech and, as a result, the acquisition of a L2 phonological system. The switch of the perceptual system to the rapid and efficient categorical mode makes L2 sounds be erroneously equated with L1 sounds at the phonetic level and subsequently connected to L1 categories at the phonological level. The issue of the L1 influence on L2

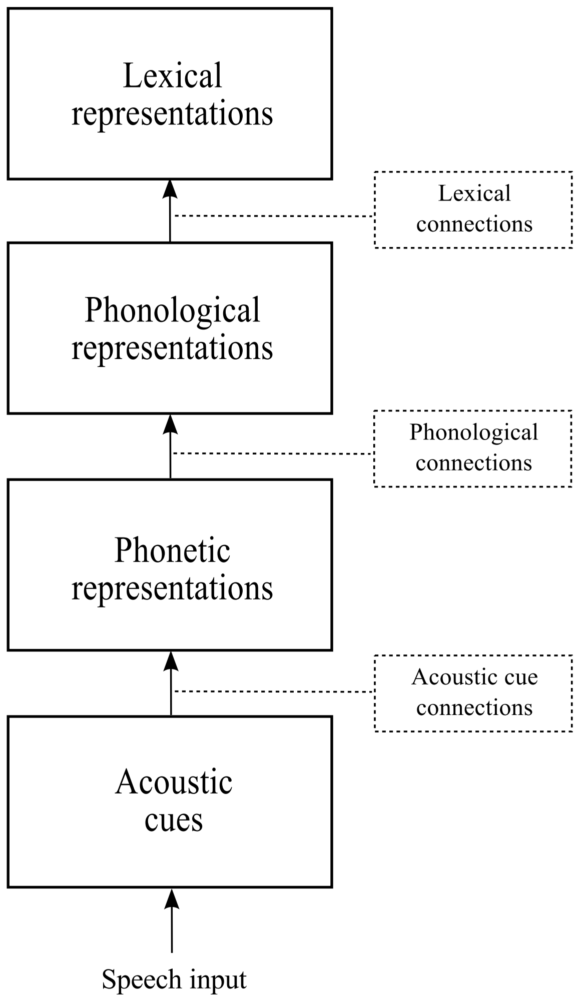


Figure 1. Model of speech perception and word comprehension adapted from Pisoni & Luce (1987), Ramus et al. (2010) and van Leussen & Escudero (2015).

speech perception and the acquisition of a L2 phonological system has been addressed by several L2 speech learning models reviewed in the following sections.

2.3 Models of L2 speech perception

Previous research on L2 speech perception has proposed several models seeking to explain and predict L2 learners' difficulties in L2 speech learning as being attributed to the influence of previous linguistic experience with the L1. The concept of perceived cross-language phonetic similarity, that is, the degree of perceived acoustic-phonetic distance between L1 and L2 sounds, is central for such models. The present section first describes the developmental changes which occur in non-native speech perception and affect learners' perception of L2 speech[1]. The current L2 speech learning models are reviewed to gain a better understanding of the mechanisms underlying L2 speech learning and the acquisition of L2 perceptual phonological competence.

2.3.1 Developmental changes in non-native speech perception

Previous research has shown that at birth infants are capable of discriminating between speech sounds not found in the L1. However, within the first year of life their ability to distinguish between non-native speech sounds rapidly declines as linguistic experience with the native language increases (Kuhl et al., 2006; Polka & Werker, 1994; Werker & Tees, 1984). Thus, the normal development of the perceptual system in the L1 involves a gradual change from a language-universal to a language-specific mode of perception. By adulthood people fail to distinguish between many non-native speech sounds. The *Native Language Magnet* model (NLM: Kuhl, 1991; Kuhl & Iverson, 1995; Kuhl et al., 2008) and the *Perceptual Assimilation Model* (PAM: Best, 1994, 1995; Best & McRoberts, 2003) have been proposed to explain the mechanisms by which linguistic experience alters discriminability of non-native speech sounds.

The NLM model explains that the developmental change from language-universal to language-specific speech perception is caused by the development of and eventual commitment to the L1 phonological system. Kuhl argues that the establishment of L1 phonetic categories is facilitated by the massive exposure to the L1 along with infants' remarkable speech processing capacity. The model holds that infants acquire L1 categories by using their innate abilities of pattern detection and statistical learning (Kuhl, 2004, also see Kuhl et al., 2008, for the expanded NLM model's predictions regarding the acquisition of two first languages simultaneously). They analyze the distributional frequency of sounds in the L1 and group the speech sounds near the *modal values* or *prototypes*. Once the L1

[1]The term non-native speech perception is used to refer to speech perception by listeners who have not been previously exposed to the target language, whereas L2 speech perception is understood as speech perception by learners of the target language.

phonetic prototypes have been formed they begin to work as *magnets* attracting nearby speech sounds on the basis of their acoustic similarity. The NLM model posits that the *perceptual magnet effect* of the L1 dramatically affects non-native speech perception and further L2 speech learning in adulthood. According to the model, L1 phonetic prototypes *warp* the acoustic space underlying speech perception by attracting acoustically similar L2 sounds and making them become perceived as instances of particular L1 phonetic categories. Due to the minimization of the within-category differences those L2 and L1 sounds, and between contrasting L2 sounds will be difficult to distinguish.

Similar to the NLM model, the Perceptual Assimilation Model holds that previous L1 exposure and the emergence of an L1 phonological system alter speech perception early in life. The PAM proposes a mechanism known as *perceptual assimilation* to account for the decline of discriminability of non-native contrasts in adulthood. The PAM follows an articulatory rather than a purely acoustic approach to explain this developmental change in L2 speech perception. Best (1994) explains that very young infants are not yet influenced by the L1 and are able to perceive articulatory and acoustic distinctions in both native and non-native sound contrasts. During the first year of life infants begin to learn articulatory patterns of L1 speech sounds. Gradually speech perception attunes to the L1 phonological system, which results in speech sounds being perceived at the level of phonological contrasts (i.e. phonological level). At this level non-native speech sounds become readily *assimilated* (i.e. perceived as equivalent) to the L1 categories on the basis of their articulatory-gestural similarities. In other words, once the L1 phonological system is established we perceive sounds in other languages as instances of our L1 sound categories.

Both the NLM and the PAM point out that the change in speech perception driven by L1 experience does not alter the sensory ability to discriminate speech sounds (Best, 1994; Kuhl & Iverson, 1995). Rather, "the change occurs at a higher level, one that involves memory and/or attention" (Kuhl & Iverson, 1995: 142). This idea has been supported by a number of studies showing a link between the decline in non-native speech perception and the development of cognitive skills in childhood (Conboy, Sommerville, & Kuhl, 2008; Diamond, Werker, & Lalonde, 1994; Lalonde & Werker, 1995; Werker & Pegg, 1992). For example, Conboy et al. (2008) found that the perception of non-native contrasts was negatively associated with inhibitory control early in life. The results of the study suggest that the decline in the ability to discriminate non-native sounds is promoted by infants' developing capacity to ignore acoustic information in speech that is irrelevant for their L1. In other words, increasing experience with the L1 and cognitive development alter speech perception in such a way that only the acoustic-phonetic details used to define phonetic categories in the native language are attended to, whereas irrelevant acoustic-phonetic information is ignored (Werker, 1995).

In summary, the experience with the L1 and the development of cognitive abilities begin to affect individuals' non-native speech perception very early in life. The observed decrease in non-native speech perception may cause difficulties for learners who acquire their L2 phonological system. This issue has been addressed by the current L2 speech learning models reviewed in the following section.

2.3.2 Explaining L2 speech learning difficulties

There are several models of L2 speech learning striving to explain learners' difficulties in acquiring L2 sounds. Similar to the models reviewed above, these models emphasize the role of previous experience with the L1 in L2 speech perception. This section reviews L2 speech learning models' predictions concerning learners' difficulties in acquiring L2 perceptual phonological competence. Previous empirical studies on L2 speech learning are also reviewed in this section. This overview aims to shed light on the perceptual mechanisms explaining learners' difficulties in the acquisition of L2 perceptual phonological competence and identify the research methods used in its assessment.

2.3.2.1 The Speech Learning Model

The *Speech Learning Model* (SLM: Flege, 1995; Flege, Schirru, & MacKay, 2003; Flege, 2007) is concerned with the difficulties experienced L2 learners have in acquiring L2 sounds, which involves the establishment of L2-specific phonetic categories. Flege (1995: 244) defines a phonetic category as a long-term memory representation for a wide range of speech sounds identified as being the same "despite auditorily detectable differences between them along dimensions that are not phonetically relevant". The SLM posits that L2 sounds become equated with L1 sound(s) due to the mechanism of *equivalence classification* that causes identification of perceptually similar L2 and L1 sounds as phonetic variants of the same L1 category and blocks formation of new categories for L2 sounds.

An important tenet of the SLM is that "the mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span" (Flege, 1995: 239). L2 learners may establish accurate L2 phonetic categories provided they have had enough L2 experience to be able to discern at least some of the acoustic-phonetic differences between perceptually close L2 and L1 sounds. The model explains L2 learners' difficulties in perceiving acoustic-phonetic differences between L2 and L1 sounds in terms of the levels of speech perception (see Section 2.2.2). Flege (1995: 241) claims that L2 learners' failure to discern the differences between L2 and L1 sounds, and between contrasting L2 sounds may occur at the acoustic and the phonological level of speech perception. On the one hand, L2 learners may initially fail to detect sensory (i.e. acoustic) differences between speech sounds pre-attentively as a result of categorical speech perception caused by the acquisition of the L1 phonological system. On the other hand, under some listening conditions, L2 learners can attend to acoustic-phonetic differences which may be discarded at the phonological level during online speech processing.

2.3.2.2 The Perceptual Assimilation Model

The *Perceptual Assimilation Model*, as applied to L2 speech learning (PAM-L2: Best & Tyler, 2007) aims at predicting experienced adult L2 learners' difficulties in learning L2 phonological contrasts in terms of perceived similarity between L2 and L1 speech sounds. In line with the SLM, the model attributes L2 learners' difficulties in L2 speech learning to the interaction between L1 and L2 sound systems and the influence of L1 phonetic categories on the perception of L2 sounds. The PAM-L2 is particularly concerned with the acquisition of L2 phonological contrasts, which is determined by the way contrasting L2 sounds are *perceptually assimilated*, that is, perceived as acoustically and phonetically similar to L1 sounds. The PAM-L2 proposes four types of perceptual assimilation patterns of L2 sounds to L1 categories that predict learners' success in acquiring L2 phonological contrasts (see Table 1).

Table 1
Four cases of L2 phonological contrast assimilation and learnability predictions.

L2 contrast cases	**Learning**	**Assimilation type**
One L2 sound = L1 category	No learning of the equivalent L2 sound	Two-Category Uncategorized-Categorized
Both L2 sounds = same L1 category, but one is perceived as a better exemplar	Successful for the "deviant" L2 sound	Category-Goodness
Both L2 sounds = same L1 category	Problematic	Single-Category
Neither L2 sound = single L1 category	Successful	Uncategorized

According to the model the learning of an L2 phonological contrast will be successful if both L2 sounds are assimilated to different L1 categories or if one of them is assimilated to an L1 category and the other one is uncategorized. The *Single-Category* and *Category-Goodness* assimilation patterns are believed to be the most problematic for L2 learners. These two assimilation patterns constitute the main focus of the present dissertation. In the case of *Single-Category* assimilation both contrasting L2 sounds are perceived as equally good or poor instances of the same L1 category. In this situation L2 learners have difficulties discriminating such L2 sounds, which are equated at the phonetic and phonological level to the same L1 category. The success in acquiring this type of L2 phonological contrast depends on the learners' ability to form a new phonetic category for at least one of the L2 sounds by detecting the acoustic-phonetic differences between L2 sounds.

In the case of *Category-Goodness* assimilation both L2 sounds are assimilated to the same L1 category, but one of them is perceived as being perceptually closer to the L1 category than the other. The PAM-L2 predicts that a new phonetic category is likely to be formed for the "deviant" L2 sound in a contrast (i.e. one perceived as a worse-fitting phonetic category of the same L1 sound). Finally, when neither of the two contrasting L2 sounds is perceived as belonging to any single L1 category, but rather to several L1 categories, new L2 phonetic categories for one or two sounds may be easily acquired.

The PAM-L2 explains L2 learners' difficulties in L2 speech learning by arguing that L2 learners shift their attention away from the phonetic level, at which the differences underlying contrasting L2 sounds are detected, and focus on the phonological level, which becomes their dominant level of processing during L2 speech perception Best and Tyler (2007).

2.3.2.3 The Second Language Linguistic Perception

The *Second Language Linguistic Perception* model (L2LP: Escudero, 2009; van Leussen & Escudero, 2015) aims at explaining the process of acquisition of L2 phonological contrasts from initial (non-native) to final (native-like) performance. Similar to the SLM and the PAM-L2, this model predicts that learners' difficulties in L2 speech learning are related to the influence of L1 categories on L2 speech perception. The L2LP, based on the optimal perception hypothesis (Escudero, 2009), states that learners initially create a copy of their L1 phonological system and L2 sounds are perceived in terms of L1 categories. Beginning L2 learners perceptually assimilate L2 sounds to L1 categories on the basis of perceived acoustic-phonetic similarity between the L1 and L2 sounds.

Over the course of L2 learning this copy is being adjusted in a L2-specific way by the means of the alternation of the existing L1 categories (i.e. shifting category boundaries or splitting categories) or the creation of new L2 phonetic categories. Similar to the PAM-L2, the L2LP model is concerned with the acquisition of L2 phonological contrasts by L2 learners. The model predicts three scenarios of learning L2 phonological contrasts L2 learners may face. In the case of the most challenging *New* scenario (the equivalent of the PAM's Single-Category assimilation type) where two L2 sounds are assimilated to a single L1 category, learners must create a new L2 category or split the existing L1 category to which both L2 sounds are assimilated. In the *Similar* scenario (the PAM's Two-Category assimilation type), in which two L2 sounds are assimilated to two L1 categories, learners must adjust L1 categories so that their boundaries match those of the L2 contrast. According to the L2LP, this scenario causes less learning problems than the New scenario. Finally, there is the *Subset* scenario (the PAM's Uncategorized or Categorized-Uncategorized assimilation type) in which one of L2 sounds in a contrast is perceived as more than one L1 category. This scenario is predicted to cause few problems since L2 learners face little difficulty in perceiving this contrast.

Similar to the SLM and the PAM-L2, this model refers to the levels of speech perception to explain how L2 phonological contrasts are learned. According to the L2LP, L2 speech learning represents an alternation of the initially created copy of the L1 phonological system. This requires attunement of the connections between different levels of speech perception in an L2-specific way. In other words, it can be said that L2 perceptual speech learning takes place if the L2 learners' perceptual system can accurately encode L2-specific acoustic-phonetic properties of L2 sounds at the phonetic level and accurately map them onto distinct L2 categories.

On the whole, learners' difficulties in acquiring L2 sounds derive from early and extensive experience with the L1, which dramatically affects the perception of L2 speech later in life. The influence of the previously acquired L1 phonological system leads to L2 learners' categorizing L2 sounds as instances of L1 categories at both phonetic and phonological levels of speech perception. The assimilation of contrasting L2 sounds to the same L1 category decreases the likelihood of acquiring such L2 phonological contrasts, that is, the likelihood of establishing L2 phonetic categories for those speech sounds. L2 speech learning may be facilitated by learners' ability to discern acoustic-phonetic differences between L2 and L1 sounds, and between contrasting L2 sounds. In other words, L2 learners' perceptual sensitivity to acoustic-phonetic properties of L1 and L2 sounds may help learners overcome the influence of the L1 phonological system on the acquisition of L2 phonological contrasts (Odlin, 1989). Taken together, the predictions of L2 speech learning models suggest that the degree of perceived acoustic-phonetic distance between L2 and L1 sounds underlying ability to establish new phonetic categories for L2 sounds is an important determinant of L2 learners' perceptual phonological competence.

This dissertation is framed within the models reviewed above. In particular, the PAM-L2 assimilation patterns are used to identify learners' difficulties in the perception of L2 vowel contrasts examined in the present study (see Chapter 5).

2.4 Perceived cross-language phonetic similarity and discrimination of L2 sounds

In order to predict and explain learners' difficulties in L2 speech learning, previous studies have assessed L2 learners' degree of perceived cross-language phonetic similarity and the ability to accurately perceive L2 phonological contrasts[2]. These abilities have often been measured through identification and discrimination tasks (see Bohn, 2002; Mora, 2008; Strange & Shafer, 2008, for comprehensive reviews). The former requires L2 learners to compare L2 sounds against internalized L1 categories and has been considered a useful method of assessing the perception of cross-language phonetic distances (Cebrian, Mora, & Aliaga-Garcia, 2011; Strange, 2007) and predicting/explaining discriminability of contrasting L2 sounds (Guion et al., 2000). The latter has been used to test learners' ability

[2]In this dissertation the terms phonetic similarity and phonetic distance are used to refer to L2 learners' perceived distance between L2 and L1 sounds, and between contrasting L2 sounds.

to establish L2 phonetic categories. This section aims at providing an overview of previous studies on L2 learners' perception of cross-language phonetic similarity and their perception of L2 phonological contrasts.

2.4.1 Degree of perceived cross-language phonetic similarity

In support of the L2 speech learning models reviewed above numerous studies have shown that perceived cross-language phonetic similarity, which may cause a pair of contrasting L2 sounds being assimilated to the same L1 sound, influences learners' success in acquiring L2 phonological contrasts (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Rallo Fabra & Romero, 2012; Flege & MacKay, 2004; Flege, MacKay, & Meador, 1999; Guion et al., 2000; Lengeris, 2009; Levy, 2009b, among others).

A study conducted by Levy (2009b), confirmed the predictions made by the PAM-L2 by showing that the Single-Category assimilation when two L2 vowel sounds are perceptually equated with the same L1 category poses a big challenge for perception of L2 vowel contrasts in both naive listeners and L2 learners. Levy examined American English listeners differing in their experience with French. She implemented a *cross-language assimilation overlap method* to analyze the perceptual assimilation data obtained via a perceptual assimilation task which required listeners to classify L2 vowel stimuli as instances of L1 vowel categories. Perceptual overlap is referred to as "... the smaller percentage of responses when two members of a pair of non-native or L2 sounds are assimilated to the same native category" (Levy, 2009: 2671). For example, in a French contrast /u/-/y/, /u/ was identified as /u/ 95% of the time and /y/ was categorized as /u/ in 6.8% of instances. Thus, the vowels perceptually overlapped at 6.8%. The results revealed significant correlations between the degree of perceptual overlap and accuracy in the discrimination of L2 vowel contrasts, suggesting that the more often two French vowels were assimilated to a single L1 category, the less accurate the discrimination of the vowel contrast was.

Similar findings were obtained by Guion et al. (2000) in a study on consonant perception. They used a *fit index*, which combined the percentage of identification and the goodness-of-fit rating, to assess the degree of phonetic similarity between L1 and L2 consonants and make predictions about the acquisition of L2 phonetic categories. Japanese listeners varying in English-language experience were asked to map English consonants onto L1 categories and provide a goodness-of-fit rating (i.e. how well the L2 sounds matched the L1 categories). The results indicated that the degree of perceived similarity between L2 and L1 sounds, as measured by the fit index, predicted the discrimination accuracy of L2 sounds. In line with PAM and PAM-L2 predictions, our study was conducted under the assumption that learners' difficulties in acquiring L2 phonological contrasts depends on the degree to which the contrasting L2 sounds are identified as instances of a single L1 category.

Consistent with Levy's (2009) and Guion et al.'s (2000) findings, research on Catalan-Spanish bilingual listeners' perception of English vowels has shown that L2 learners' difficulties in learning L2 phonological contrasts are related to learners' assimilating L2 vowels to L1 vowel categories (Cebrian, 2007; Rallo Fabra & Romero, 2012). For example, Rallo Fabra and Romero (2012) tested the ability of Catalan learners of English to discriminate seven Catalan–English (/a/–/æ/, /a/–/ɑ/, /ɛ/–/ɛ/, /a/–/ʌ/, /i/–/i/, /i/–/ɪ/, and /u/–/u/) and four English–English (/ɑ/–/ʌ/, /ɛ/–/æ/, /i/–/ɪ/, and /u/–/ʊ/) vowel contrasts. The results showed that Catalan learners had difficulties in perceiving the Catalan-English /a/–/æ/, /ɛ/–/ɛ/, and /a/–/ʌ/ vowel pairs and English /ɑ/–/ʌ/ and /ɛ/–/æ/ vowel pairs. The observed discrimination difficulties were explained by the L2 learners' tendency to assimilate perceptually similar L2 and L1 sounds and L2 vowel pairs to the same L1 category.

Overall, the findings of previous studies confirm that L2 learners' difficulties in acquiring L2 phonological contrasts are attributed to the differences between L2 and L1 phonological systems and degree of perceived cross-language phonetic similarity which make L2 learners' categorize L2 sounds as instances of the same L1 category. These factors affect L2 learners' accuracy in discriminating such L2 sounds and impede the acquisition of L2 phonological contrasts. Taken together, previous research indicates the importance of measuring L2 learners' degree of perceived phonetic distance between L2 and L1 speech sounds as well as their ability to discriminate contrasting L2 sounds in assessing L2 learners' speech perception.

2.4.2 Discrimination of contrasting L2 sounds

Whereas perceptual assimilation tasks have been primarily used in previous speech perception research to predict and explain L2 learners' difficulties in acquiring L2 sound contrasts, the ability to discriminate contrasting L2 sounds has long been used as a measure of L2 perceptual phonological competence (Darcy et al., 2014, 2015; Gottfried, 1984; Højen & Flege, 2006). The assessment of this ability requires taking into account acoustic cue weighting and task design.

2.4.2.1 Weighting of L2 acoustic cues

Acoustic cue weighting refers to the differential contribution of some acoustic cues (e.g., formant frequency, duration and voice onset time) to the categorization of speech sounds (Holt & Lotto, 2006). L2 learners' use of acoustic cues in the discrimination of contrasting L2 sounds has received much attention in research on L2 speech perception (Bohn, 2002; Cerviño-Povedano & Mora, 2011; Flege et al., 1997; Holt & Lotto, 2006; Lengeris, 2009; McAllister, Flege, & Piske, 2002; Ylinen et al., 2010). There exist two views on learners' use of L2 acoustic cues while perceiving L2 speech sound contrasts. On the one hand, some researchers have argued that L2 acoustic cue weighting is highly dependent on learners' L1 background and L2 experience, that is, it is subject to linguistic experience (Flege et al., 1997; McAllister et al., 2002). For example, McAllister et al. (2002) examined the production and perception of the Swedish vowel length contrast by native speakers of Swedish, Estonian, English and Spanish, which differ in the use they make of vowel duration. The results showed that L1-Estonian speakers, who use temporal cues to distinguish vowels in their L1, were more target-like in the perception and production of Swedish vowel contrasts than L1-English and L1-Spanish speakers who do not use duration to distinguish vowels in their L1. McAllister et al. proposed the *Feature hypothesis*, which posits that L2 features like temporal cues not used in the learners' L1 to signal phonological contrasts will be difficult to perceive and produce.

On the other hand, several studies have shown that the use of temporal cues may be relatively language independent (Bohn, 1995; Bohn & Flege, 1990; Cebrian, 2006; Cerviño-Povedano & Mora, 2011). For example, investigating the perception of American English vowels by native Spanish and German learners, Bohn (1995, Bohn & Flege, 1990) found that adult L2 learners tended to rely more heavily on duration rather than on spectral cues when identifying English vowels[3]. Taking into account the fact that duration is used phonologically in German but not in Spanish, Bohn proposed the *Desensitization hypothesis* which states that "... duration cues in vowel perception are easy to access whether or not listeners have had specific linguistic experience with them" (Bohn, 1995: 294). Similarly, Lengeris (2009) suggested that duration does not have a special status in L2 vowel perception and may be available to L2 learners who do not use duration in their L1.

Both Bohn (1995) and McAllister et al. (2002) point out that the extent to which temporal cues are used in the discrimination of L2 speech sounds may depend on the task procedure, as in a two-alternative forced-choice test. The role of the task design in the assessment of L2 learners' ability to discriminate L2 sounds is discussed below.

2.4.2.2 Effects of task variables on discrimination

The effects of task variables, such as the nature of the stimuli and the inter-stimulus interval, on the discrimination of speech sounds has long been of concern in speech perception research (see Bohn, 2002; Mora, 2008; Strange & Shafer, 2008, for discussion). Several task procedures have been adopted in L2 speech research for assessing L2 learners'

[3]Both spectral (formant frequency) and temporal (duration) acoustic cues differentiate English tense and lax vowels like /i/ and /ɪ/. However, native speakers of English rely much more on the spectral dimension than the temporal dimension in categorizing tense-lax vowels (Holt & Lotto, 2006).

discrimination ability (see Flege, 2003; Mora, 2008, for a review of existing methods). Among them are AX, AXB, ABX and oddity discrimination tasks. The common feature of these tasks is that they aim at measuring whether L2 learners have established new phonetic categories for L2 sounds. In order to assess the establishment of L2 phonetic categories, previous studies have widely used a categorical discrimination task. This task measures learners' ability to accurately discriminate L2 sounds at the phonological rather than the acoustic level of speech perception, which is controlled by varying the interstimulus interval or using multiple speakers' voices.

For example, Gottfried (1984) designed a cross-speaker categorical ABX discrimination task to examine the perception of French vowels by English inexperienced listeners and L2 learners of French. The participants were presented with a triad of monosyllabic stimuli produced by three different speakers. The vowel stimuli A and B were tokens of a different vowel category and X was a token of the categories of either A or B. Thus, the task required listeners to categorize vowels as one vowel category or the other while ignoring inter-speaker variation. A 1000 ms interstimulus interval was used to increase the memory load. Using this task Gottfried showed that both experienced and inexperienced L2 listeners had difficulty in categorically perceiving L2 vowel contrasts.

More recent studies, extending Gottfried's task, have further increased task demands by using more complex stimuli in which the target vowels are embedded in disyllabic (Levy & Strange, 2008) or trisyllabic (Darcy et al., 2014) nonwords. Thus, a cross-speaker categorical discrimination task based on vowels presented in a more complex context and produced by several speakers more closely resembles what listeners do during speech processing. As Strange and Shafer (2008: 167) pointed out, such type of tasks can be considered "a sensitive measure of differences in higher-order phonetic perception processes in L1 and L2 listeners, and may be a better measure of their perceptual capabilities in real world situations".

The importance of using a categorical discrimination task was demonstrated by Højen and Flege (2006) who tested Spanish monolinguals' perception of English vowel contrasts using a categorical AXB task. In their task a manipulated fundamental frequency (F0) of the stimuli and interstimulus intervals of 0 ms and 1000 were used to avoid the participants' reliance on auditorily detectable differences between the speech stimuli. The results showed that in comparison with native English speakers early Spanish learners of English obtained near-chance scores for three difficult vowel contrasts (/ɪ/-/eɪ/, /a/-/ʌ/ and /ʊ/-/oʊ/), suggesting that participants had difficulties in perceiving these English vowels as distinct vowel categories. It is important to mention that the categorical task used by Højen and Flege (2006) hindered lower acoustic level of speech perception. Specifically, the task forced the participants to perceive sound contrasts at the phonological level which operates with the existing categories and did not let them rely on the acoustic differences between the speech stimuli.

Overall, the results of the studies reviewed above suggest that in the assessment of L2 learners' discrimination of contrasting L2 sounds it is important to be aware of the task factors which may affect L2 learners' performance. Using multiple speakers' voices and rather long interstimulus intervals may help ensure that the higher phonological level rather than the lower acoustic level of speech perception is assessed. In other words, the task measuring discrimination of L2 sounds should resemble what listeners do during natural speech processing.

2.4.3 Cross-language similarity of English vowels to Catalan vowels

In this dissertation we examine the perception of five General American English vowel contrasts /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ in Catalan learners of English. These particular vowel contrasts were chosen on the basis of previous L2 speech perception research showing that Catalan learners of English may have difficulties in perceiving these vowel contrasts. These difficulties have been attributed to differences between the English and Catalan vowel inventories and the perceived phonetic similarity between Catalan and English vowel sounds (Cebrian, 2006; Cebrian et al., 2011;

Cerviño-Povedano & Mora 2011; Mora & Fullana, 2007; Rallo Fabra, 2005; Rallo Fabra & Romero, 2012).

Catalan has a smaller vowel inventory than General American English (see Figure 2), with three front vowels, /i/, /e/, and /ɛ/, three back vowels, /ɔ/, /o/, and /u/) and two central vowels, /ə/ and /a/ (Rallo Fabra & Romero, 2012). According to the cross-language data provided by previous studies (see Cebrian, 2002; Rallo Fabra & Romero, 2012), the English vowels in the present study are located in close proximity of Catalan vowels.

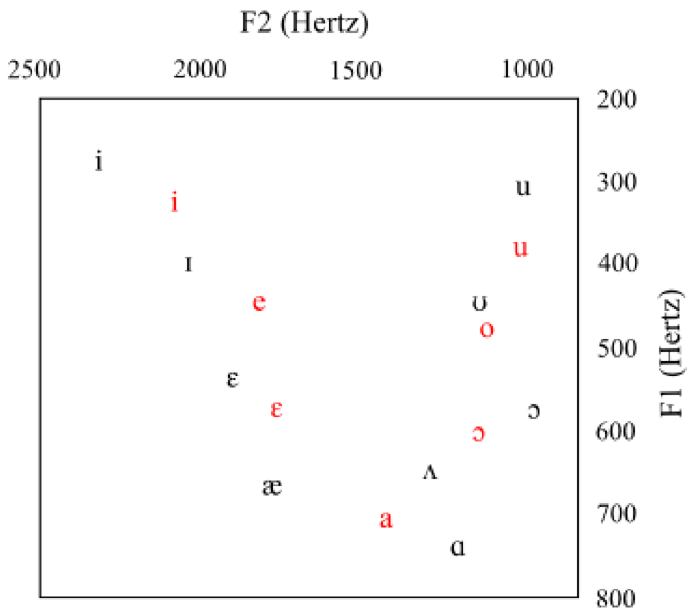


Figure 2. Vowel charts for Eastern Catalan (red) and General American English (black). The spectral frequency values are borrowed from Rallo Fabra & Romero (2012: 493-494).

For example, English /i/ and /ɪ/ are acoustically and perceptually close to Catalan /i/. English /ɪ/ is close to Catalan /e/ and English /ɛ/ may be perceived as intermediate between Catalan /e/ and /ɛ/. English /æ/ occupies an intermediate position in the acoustic space between Catalan /ɛ/ and /a/. The low central and back English vowels /ʌ/ and /ɑ/ are located in the proximity of Catalan /a/. Vowel inventory data along with PAM-L2 predictions (Best & Tyler, 2007) suggest that Catalan learners of English may have difficulties in acquiring the L2 vowel contrasts we target in the present study due to single-category assimilation of contrasting L2 vowels to perceptually close L1 vowel categories.

2.5 Summary

This chapter has set the theoretical framework for the part of this dissertation concerning the acquisition of L2 perceptual phonological competence. The chapter has discussed the most important issues in L2 speech learning that have been emphasized by previous speech perception research. As can be inferred from previous research outcomes, early changes in L2 speech perception, caused by the already established L1 phonological system, significantly affect subsequent learning of L2 phonological contrasts. L2 learners' difficulties in acquiring L2 phonological contrasts are attributed to learners' tendency to map L2 sounds onto a single perceptually similar L1 sound category.

In order for successful L2 speech learning to take place, L2 learners must be able to establish accurate L2 phonetic categories which will be further mapped onto distinct phonemes. The establishment of L2 phonetic categories is determined by the ability to discern acoustic-phonetic differences between L2 and L1 sounds, and between contrasting L2 sounds. The ability to acquire new L2 categories might require L2 learners to attend to both continuous and categorical sources of information in L2 speech input. That is, during L2 speech perception both acoustic and phonological sources of information will contribute to learners' perceptual sensitivity to acoustic-phonetic properties of L2 sounds and help them shape their L2 phonological system.

The evidence provided by the studies reviewed in this chapter suggests that L2 learners' success in acquiring L2 sounds is facilitated by learners' ability to perceive a cross-language phonetic distance between L2 and L1 sounds, and is indicated by their ability to accurately discriminate contrasting L2 sounds. The former can be measured using a perceptual assimilation task which requires identification and similarity judgment of L2 speech sounds as instances of L1 categories. The latter can be assessed through a categorical discrimination task which requires L2 learners to discriminate L2 sounds in a categorical manner. These tasks are used in our study to assess L2 learners' perception of L2 vowel sounds.

As discussed in Chapter 1, several factors have been shown to contribute to the acquisition of both perceptual and productive phonological competence. Among those factors are previous experience with the L2, general level of competence in the L2, and the amount and quality of the L2 input received, all of which have been thoroughly investigated in previous SLA research. However, cognitive mechanisms involved in L2 speech perception may also contribute substantially to learners' success in L2 speech learning and still remain under-researched. The next chapter reviews this line of research.

Chapter 3

Cognitive ability in L2 speech learning

3.1 Introduction

It is widely observed that individuals differ in their ability to perform certain tasks. For example, some people are much better at drawing or at story-writing than others. This may be attributed to individual differences, that is, "dimensions of enduring characteristics that are assumed to apply to everybody and on which people differ by degree" (Dörnyei, 2005: 4). Individual differences have long been observed and extensively investigated in the field of SLA (see Dörnyei, 2005; Segalowitz, 1997, for a review of individual differences), but have received little attention in research on L2 phonological acquisition.

Regarding inter-learner variation in L2 phonological competence, there exist exceptionally talented L2 learners who manage to acquire native-like competence in L2 speech perception and production despite learning the language quite late in life (Birdsong, 2007; Bongaerts, 1999; Ioup et al., 1994; Moyer, 2014). On the other hand, many L2 learners struggle with L2 pronunciation and find it very hard to attain a high level of L2 perceptual (and productive) phonological competence. Over the last decade, studies investigating inter-learner variation in L2 phonological competence have taken a psycholinguistic approach investigating the relationship between L2 learners' individual differences in cognitive ability and L2 learning. The individual differences in cognitive ability have been put forward as candidates to explain inter-learner variation in L2 phonological acquisition in this line of SLA research.

Research on executive functions has identified three core cognitive abilities that are essential for the performance of complex cognitive tasks, such as problem solving, decision making and learning. They are working memory (implies storing information in mind and working with it), inhibition (disactivation of irrelevant information) and attention shifting (flexible switching between the demands of different tasks), with the latter two being important aspects of attention (or executive) control (Baddeley, 2002a; Diamond, 2013; Miyake et al., 2000). Attention control can be described as an umbrella concept covering a variety of mechanisms that underlie our capacity to choose what we focus attention on and what we ignore (Astle & Scerif, 2009). Working memory and attention control are multicomponent and interrelated abilities which support and need one another (Diamond, 2013). Specifically, we focus on cognitive abilities such as phonological short-term memory (a working memory component in charge of temporarily storing verbal information), acoustic short-term memory (a component of working memory in charge of temporarily holding acoustic properties of speech sounds), and attention-shifting function of attention control.

The present chapter provides an overview of previous research on cognitive abilities involved in information processing and language learning. The current model of working memory is provided. The chapter also reviews previous research on working memory and attention control as regards their role in L2 speech learning. This will help understand how cognitive ability contributes to the acquisition of L2 perceptual phonological competence.

3.2 Memory and attention in information processing and L2 learning

Previous research investigating aptitude as a factor in SLA has focused on individual differences in cognitive ability as a means of explaining inter-learner variation in L2 learning. L2 learning aptitude may be defined as "... strengths individual learners have—relative to their population—in the cognitive abilities information processing draws on during L2 learning..." (Robinson, 2005: 46). L2 learning aptitude is thus viewed as a multicomponent construct consisting of cognitive abilities involved in L2 learning (Dörnyei & Skehan, 2003; Robinson, 2005, 2012; Skehan, 2002, 2012). Working memory has been proposed as a central component of L2 learning aptitude (Dogil & Reiterer, 2009; Jilka et al., 2008; Miyake & Friedman, 1998; Skehan, 1998, 2012). The purpose of the present section is to introduce the concept of working memory and determine its role in language learning.

3.2.1 The working memory model

Information processing can be referred to as manipulation of new information. It is involved, for instance, in learning a poem by heart or making a decision about which car to buy. It comprises encoding, storing, retrieving and transformation of information (see Figure 3) and takes place in three stages (Atkinson & Shiffrin, 1968; Baddeley, 1997, 1999; Wickens, Gordon, & Liu, 1997). It starts with initial encoding of information coming from our senses (sight, hearing, taste, smell and touch). The information that receives attention is further transferred to working memory for further processing. Working memory can be compared to a *mental workbench* where information from sensory memory and long-term memory is temporarily maintained so that it can be evaluated, transformed and acted on (Baddeley, 2012, 2002b; Wickens et al., 1997). Finally, the processed information can be encoded into the unlimited and long-lasting long-term memory, which passively stores information for later use. This memory storage can be thought of as a repository of our acquired knowledge. Thus, working memory takes a central role in transferring information to long-term memory, which underlies the acquisition of new knowledge.

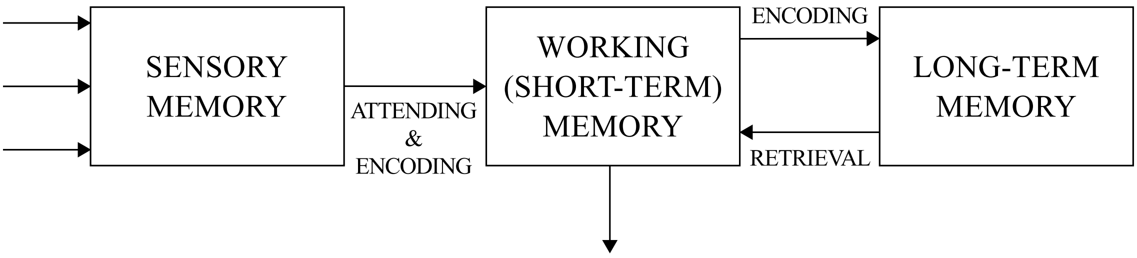


Figure 3. The information processing model adapted from Atkinson & Shiffrin (1968) and Wickens et al. (1997).

According to the influential model (see Figure 4) proposed by Alan Baddeley (2002b, 2012, 2000) working memory is a multicomponent system that consists of three short-term memory stores responsible for temporarily holding different types of information and attention-related processes, termed the central executive (Cowan, 2008; Baddeley, 2000, 2012, also see Miyake & Shah, 1999, for alternative working memory models). The visuo-spatial sketchpad is a short-term memory store which holds and manipulates information about objects and locations. The phonological loop is a short term-memory storage mechanism responsible for temporarily maintaining verbal-acoustic information and its serial order in memory. The episodic buffer is in charge of the temporary storage of integrated episodes from multiple dimensions. It connects the so-called slave systems (the phonological loop and the visuo-spatial sketchpad) to the long-term memory store. Finally, the central executive is a multifunctional system responsible for the supervision of the slave subsystems, attention control (focusing, dividing and switching attention) and activation of long-term memory (i.e. retrieving information from and encoding it to long-term memory (Baddeley, 1996b, 2002b, 2012). Working memory is considered crucial for learning new information, and for comprehension and decision-making (Baddeley, 1996a, 2012). In the next section we review previous research on the role of working memory in language learning.

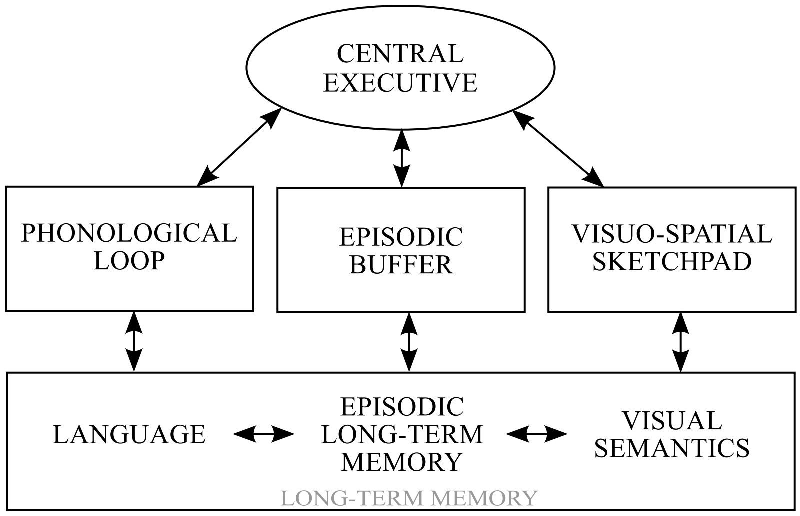


Figure 4. The multicomponent model of working memory adapted from Baddeley (2000, 2012).

3.2.2 Working memory in language learning

The short-term storage and attentional components of the working memory model described above have been considered essential cognitive abilities implicated in successful language learning (Robinson, 2003; Skehan, 1998, 2012, 2002). In order for new linguistic information to be learned, it must be processed. As discussed above, this requires relevant information to be attended to, manipulated in working memory and, finally, encoded into long-term memory. Attentional mechanisms such as focusing, inhibition, detection and attention-shifting are critical to language processing. They are responsible for encoding linguistic material in working memory, keeping it active, and for retrieving existing knowledge from and transferring new knowledge to long-term memory (Robinson, 2003; Segalowitz, 2010; Tomlin & Villa, 1994).

These cognitive abilities are thought to be implicated at various L2 learning stages (Skehan, 1998, 2002, 2012). The stages are segmenting the linguistic input stream, detecting relevant information and inferring new language rules and modifying existing ones on the basis of detected patterns. The ability to hold information temporarily in short-term memory and drawing attention to its properties play a crucial role in the process of acquiring new representations and restructuring existing ones held in long term-memory.

The stages in learning L2 sounds may be identified as following: segmenting the speech stream, detecting relevant acoustic-phonetic properties that distinguish L2 sounds and establishing new L2 phonetic categories. Thus, the short-term storage for L2 speech and attentional mechanisms may potentially play an important role in L2 speech learning. Short-term memory and attention control capacity are limited and vary among individuals (Cowan, 2001, 2005). Consequently, because short-term memory and attention control are involved in L2 learning, individual differences in these cognitive abilities can be a source of individual differences in L2 acquisition.

The following sections discuss the role of individual differences in short-term memory capacity and attention control in determining L2 learners' perceptual phonological competence. We focus on the phonological loop and the central executive components of Baddeley's model of working memory described above.

3.3 Short-term memory stores for speech sounds

As discussed in Chapter 2, the perceptual system simultaneously operates in a continuous mode, processing the acoustic properties of the speech sounds in the incoming input, and in a categorical mode, assigning these speech sounds to phonological representations stored in long-term memory. Research on speech perception suggests the existence of two short-term memory stores for the categorical and continuous information about the speech input (Fujisaki & Kawashima, 1970; Joseph et al., 2015; Pisoni, 1973; Tanaka & Nakamura, 2004). One of them is thought to be responsible for the temporary storage of phonological (i.e. categorical) information, whereas the other temporarily holds acoustic (i.e continuous) properties of speech sounds. In this dissertation these short-term memory stores are referred to as phonological short-term memory and acoustic short-term memory. What is the place of these short-term stores in working memory? How do these short-term memory stores contribute to L2 speech learning? This section aims to answer these questions by reviewing previous research on the phonological loop component of Baddeley's working memory model (Baddeley et al., 1998).

3.3.1 Phonological short-term memory

3.3.1.1 Phonological short-term memory capacity

The phonological short-term memory associated with Baddeley's phonological loop component of working memory is a limited capacity store responsible for the temporary maintenance of auditory verbal (i.e. phonological) information (Baddeley, 1992, 2007, 2012; Baddeley et al., 1998). As shown in Figure 5, the phonological loop consists of two main components, a phonological short-term store that temporarily holds phonological information and a sub-vocal rehearsal mechanism that allows refreshing of the decaying phonological information by the means of inner vocalization (Baddeley, 2000; Tanaka & Nakamura, 2004). Its capacity is said to vary among individuals, but on average it is capable of holding the amount of material one can produce within two seconds (Baddeley, 1992). According to Baddeley, the phonological loop mainly operates at the phonological level. More specifically, the phonological short-term memory holds phoneme- or syllable-sized phonological units (Baddeley, 1992; Baddeley et al., 1998).

A typical example of phonological short-term memory at work is memorizing a telephone number, doing arithmetical calculations like 2+5*8, or comprehending this very sentence. In order to do these simple tasks, one has to mentally hold and repeat the sequence of numbers or words and retrieve them in a correct order without paying much attention to details like the size or font of the numbers or text. Thus, the function of the phonological short-term memory is to temporarily hold phonological information while the long-term phonological representations are being formed (Baddeley et al., 1998). This function of phonological short-term memory has been investigated in relation to language learning (see Baddeley, 2003; Gathercole & Thorn, 1998, for a review).

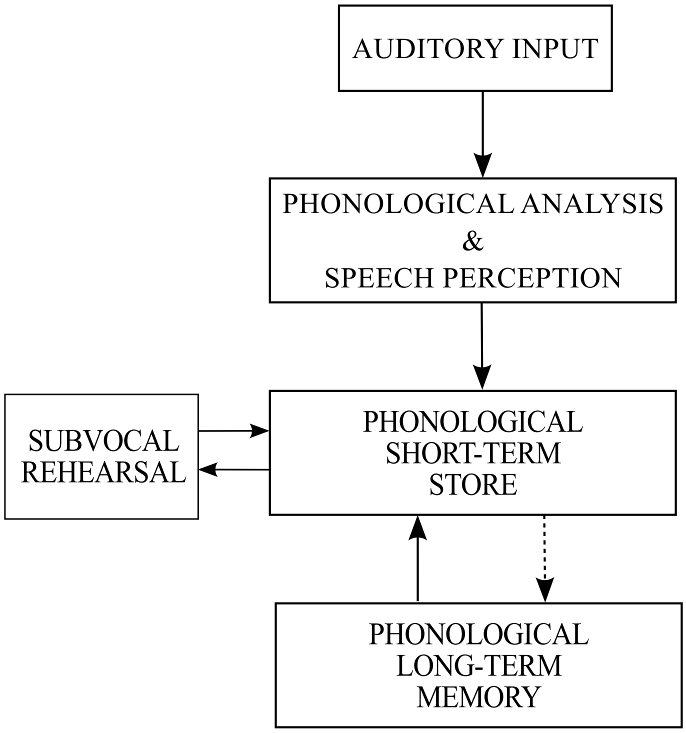


Figure 5. The phonological loop model (Baddeley et al., 1998, Baddeley, 2003).

3.3.1.2 Phonological short-term memory in L2 speech learning

Individuals' capacity to hold phonological information and its serial order in phonological short-term memory has long attracted the attention of researchers in the field of L1 and L2 language acquisition. Phonological short-term memory has been considered a "language learning device" and one of the most important cognitive abilities contributing to L2 learning aptitude (Baddeley et al., 1998; Miyake & Friedman, 1998). In language acquisition research phonological short-term memory has been found to be a strong predictor of vocabulary learning in L1 by children (Gathercole, Willis, Emslie, & Baddeley, 1992; Gathercole, Service, Hitch, Adams, & Martin, 1999) as well as in L2 by children and adults (Atkins & Baddeley, 1998; Masoura & Gathercole, 1999; Papagno & Vallar, 1995; Service, 1992). Several SLA studies have associated phonological short-term memory with acquisition of L2 grammar (French & O'Brien, 2008; O'Brien et al., 2006) and L2 phonology (Cerviño-Povedano & Mora, 2011, 2015; MacKay et al., 2001; Mora & Safronova, submitted; Tanaka & Nakamura, 2004).

Regarding L2 speech learning, MacKay et al. (2001), for instance, have provided evidence of the contribution of phonological short-term memory to inter-learner variation in L2 speech learning. They examined the perception of English consonants by 72 L1-Italian learners of English. Their phonological short-term memory was assessed through a task that required learners to repeat Italian nonwords of increasing syllable length. The phonological short-term memory was negatively correlated with participants' error rate in consonant identification, explaining 15% of the unique variance in the subjects' identification of word-final consonants, and 8% of the variance in word-initial consonants. The researchers suggested that phonological short-term memory could play an important role in L2 speech perception by facilitating the establishment of L2 phonetic categories.

Recently, studies conducted by Cerviño-Povedano and Mora (2011, 2015) have provided evidence for the role of phonological short-term memory in L2 speech perception and phonetic cue weighting. These studies on Catalan-Spanish bilingual EFL learners' perception of English vowel contrasts have demonstrated that phonological short-term memory, as measured through a serial nonword recognition task, is associated with target-like perception of L2 vowels. The results of both studies showed that L2 learners with greater phonological short-term memory capacity used a more native-like cue weighting in L2 vowel perception (i.e. could better attend to spectral cues) than the participants with lower phonological short-term memory capacity.

To sum up, previous SLA research has shown that individuals' phonological short-term memory capacity may play an important role in L1 and L2 acquisition and L2 speech learning, in particular. As discussed in Chapter 2, in order to establish new phonetic categories for L2 sounds learners need to be able to perceive acoustic-phonetic properties of L2 sounds (Flege, 1995; Flege at al., 2003). The analysis of L2 speech input for such information may not only be aided by the ability to hold verbal information in phonological short-term memory but also by acoustic short-term memory.

3.3.2 Acoustic short-term memory

Although the phonological loop has been considered responsible for storing acoustic verbal information primarily at the phonological level, it may temporarily store the acoustic properties of speech and non-speech sounds (Baddeley, 2007; Friedrich, 1990). Recently, Williamson et al. (2010) have suggested that the original working memory model may need to be improved by including an additional short-term store for detailed acoustic information such as pitch in music and speech. They argue that the acoustic and phonological stores are fundamentally different, and thus do not overlap but share a common rehearsal mechanism. This calls for research on the nature of the acoustic store and its potential role in L2 speech learning.

3.3.2.1 Acoustic short-term memory capacity

The acoustic short-term memory, as viewed by early research on speech and memory is a temporary storage for acoustic properties of sounds encoded at the acoustic level of speech perception (see Chapter 2), that is, prior to the level of phonological encoding (Cowan & Morse, 1986; Crowder & Morton, 1969; Darwin & Baddeley, 1974; Frankish, 2008; Friedrich, 1990). Acoustic short-term memory is believed to temporarily hold relatively raw acoustic information which rapidly decays after a period of 200-350 ms (Baddeley, 1997; Cowan, 1984). Similar to phonological short-term memory, acoustic short-term memory comprises a subvocal rehearsal mechanism that allows acoustic information (e.g., pitch) to be temporarily maintained in working memory, which facilitates accuracy in the discrimination of speech sounds (Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Keller & Saults, 1995; Koelsch et al., 2009).

Acoustic short-term memory capacity has been previously assessed using non-speech stimuli in tasks requiring subjects to recall auditorily presented stimuli (Li, Cowan, & Saults, 2013; Prosser, 1995). It has been found to be limited in capacity to two or three items and to depend on the quality of the items to be recalled. For example, Li et al.

(2013) estimated the acoustic short-term memory capacity using a tone probe task and a sequence recognition task. The former required participants to recall the serial position of a tone probe in the target sequence (increasing from two to six items). In the latter two sequences of tones were presented, with or without a change in one of the tones and the participants had to decide if the sequences were the same or different. They found the acoustic short-term memory capacity to be three items as measured by the single tone probe task and two items as measured by the sequence recognition task. They also found that when using multidimensional sounds (by adding timbre from musical instruments) instead of pure tones as stimuli in the single tone probe task, the acoustic short-term memory capacity increased. They explained that the acoustic short-term memory capacity increases with increasing acoustic dissimilarity between sounds.

3.3.2.2 Acoustic short-term memory in L2 speech learning

Previous speech perception research has shown that acoustic short-term memory is involved in speech perception and, in particular, in the discrimination of vowels and consonants, because it stores acoustic information necessary to distinguish speech sounds (Darwin & Baddeley, 1974; Fujisaki & Kawashima, 1970; Joseph et al., 2015; Pisoni, 1973, 1975). For example, acoustic short-term memory is believed to play a crucial role in the discrimination of contrasting speech sounds in an ABX discrimination task (Fujisaki & Kawashima, 1970). In this task participants hear a triad of speech sounds and have to decide which of the first two sounds matches the third. If the A and B (e.g., English /ɑ/ and /ʌ/) sounds are identified as being the same phonetic categories (e.g., /ɑ/ and /ʌ/ identified as single native /a/ vowel by a Russian speaker), the decision on X has to be made on the basis of acoustic details stored in acoustic short-term memory.

A recent study by Joseph et al. (2015) has demonstrated that acoustic short-term memory determines accuracy in the perception of speech sounds. In their experiment participants memorized sequences of speech sounds, containing one, two, or four syllables and had to recall one of the speech sounds in the sequence. They found that acoustic short-term memory was useful when participants were presented with the one syllable sequences. When the number of items to be held in memory increased accurate perception of speech sounds decreased.

The contribution of acoustic short-term memory to learners' perception of L2 sounds has received little attention in SLA. To the best of our knowledge, only a few studies have investigated the role of acoustic short-term memory in L2 speech perception and production (Mora & Safronova, submitted; Safronova & Mora, 2012a, 2012b; Tanaka & Nakamura, 2004). For instance, Tanaka and Nakamura (2004) investigated the role of acoustic short-term memory in the L2 pronunciation of 30 adult Japanese learners of English. Acoustic short-term memory was measured through a discrimination task that presented participants with pairs of sequences of pure tones that increased in length from two to six sounds. The subjects had to decide if the two sequences in a pair were the same or different in terms of either melody or pitch. The acoustic short-term memory scores were then related to native speakers' ratings of the learners' accuracy and fluency of pronunciation in English (segmental features, stress position, intonation and rhythm). The results showed that acoustic short-term memory (specifically, short-term memory for melody) was positively correlated with pronunciation ratings of both segmental and prosodic features.

There is evidence that the perception of speech and non-speech may be different in nature and that L2 speech learning may have a speech-specific origin (Díaz, Baus, Escera, Costa, & Sebastián-Gallés, 2008; Surprenant & Watson, 2001). Recent studies on acoustic memory have used speech-like stimuli to measure L2 learners' acoustic short-term memory (Mora & Safronova, submitted; Safronova & Mora, 2012a, 2012b). Mora and Safronova (submitted) and Safronova and Mora (2012a, 2012b) measured subjects' acoustic short-term memory capacity using rotated speech which preserves speech-specific acoustic details of normal speech (i.e. pitch changes, spectral and temporal information) but lacks intelligibility since it cannot be phonologically encoded (Scott, Blank, Rosen, & Wise, 2000; Scott, Rosen, Lang, & Wise, 2006). Using a serial nonword recognition task procedure they assessed the role of acoustic short-term memory in Catalan-Spanish learners' cue weighting in L2 vowel perception. The results indicated that L2 learners with larger acoustic short-term memory capacity are better able to rely on the spectral differences underlying the English tense-lax /i/-/ɪ/ vowel contrast. The authors suggested that a larger acoustic short-term memory capacity facilitates target-like L2 cue weighting

Although acoustic short-term memory has received little attention in SLA research, the existing evidence suggests that this memory store for acoustic information may contribute to L2 learners' perception of acoustic-phonetic details of L2 sounds. As discussed in Chapter 2, this ability is crucial for learners' acquisition of L2 perceptual phonological competence. Thus, more research is needed to investigate the relationship between acoustic short-term memory and L2 speech learning. Moreover, learning L2 sounds may require not only storage of sequences of speech sounds and their acoustic properties but also focusing attention on the relevant phonological and acoustic information during speech processing. The role of the attention control is discussed below.

3.4 Attention control

3.4.1 Attention control functions: attention shifting

Attention (or executive) control, one of the core human executive functions, constitutes one of the responsibilities of the central executive component in Baddeley's working memory model (see Figure 2). As Baddeley (1996b: 8) points out: "One important role of the central executive should be to act as an attentional controller, selecting certain streams of incoming information and rejecting others". Research on the functions of central executive has suggested that it is not a unitary system, rather it should be viewed as a complex multi-functional system that can be divided into several attention control functions such as focusing attention, dividing attention, selective attention, inhibition and attention shifting (Baddeley, 1996b, 2002a; Miyake et al., 2000).

Attention shifting, which is the focus of this dissertation, also referred to as attention switching or task switching, is considered an important aspect of executive control (Miyake et al., 2000). Its function concerns shifting back and forth between multiple tasks and involves inhibition of the information irrelevant to the performance of the current task (Diamond, 2013; Miyake et al., 2000; Monsell, 2003). To switch from one task to another, we need to inhibit (or deactivate) previous task and load a new task into working memory (i.e. activate). In this sense attention shifting requires inhibition (Diamond, 2013; Miyake et al., 2000). Previous research on attention shifting has used a dual task-switching procedure which requires shifting attention between two task sets, such as deciding on the digit (odd or even) and the letter (vowel or consonant), or the color and the shape of figures (Jersild, 1927; Rogers & Monsell, 1995; Rubin & Meiran, 2005; Prior & MacWhinney, 2010, also see Monsell, 2003, for a review of task-switching paradigms).

A common outcome of the attention control tasks is that individuals' performance is faster and more accurate on repeat than on shift trials, even if the alternation between task sets is predictable. Therefore, two basic measures of the attention-shifting skill have been used, namely, shift costs, calculated as the difference between the reaction times in shift and repeat trials, and error rates (Monsell 2003; Prior & MacWhinney, 2010; Rogers & Monsell, 1995). Error rates reflect the number of incorrect responses and serve as a measure of test takers' ability to stay on focus for repeat trials and to refocus their attention for shift trials. These measures are thought to indicate the load placed on an individual's processing system when having to refocus attention (Segalowitz & Frenkiel-Fishman, 2005).

Attention-shifting skill is important for the accomplishment of many simple every-day situations such as answering the phone while writing an e-mail. While completing these simple tasks one must shift focus of attention to select and do those which are relevant to the current goals, while suppressing irrelevant tasks (Monsell, 2003). Attention shifting is also involved in the performance of other cognitively complex tasks such as language learning, which is discussed in the following section.

3.4.2 Attention shifting in L2 speech learning

Learners' acquisition of L2 competence requires rapid and flexible attention control, which implies foregrounding relevant and backgrounding irrelevant information when decoding linguistic input. This has been demonstrated by Segalowitz and Frenkiel-Fishman (2005) who investigated the relationship between individual differences in English/French bilinguals' attention-shifting ability and L2 proficiency. They used Rogers and Monsell's (1995) alternating runs task-switching procedure to present the participants with two linguistic tasks: time adverbials (e.g. *now* vs. *later*) and causal connectives (e.g. *because* vs. *despite*). The participants had to shift between the two tasks alternating predictably in a sequence "...time-time cause-cause time-time...". The shift costs, that is, the difference between reaction times on shift (time-cause, cause-time) and repeat (time-time, cause-cause) trials and error rates were used as measures of attention control. The results showed that bilinguals who were better able to shift their attention efficiently were more proficient in their L2 (French). Moreover, attention-shifting ability was found to account for 32% of the unique variance in L2 proficiency. These results may suggest that efficient attention-shifting ability plays an important role in attaining high levels of L2 competence.

Attention control has also been shown to play an important role in L2 speech learning (Francis et al., 2000; Francis & Nusbaum, 2002; Guion & Pederson, 2007). Francis et al. (2000), for example, found that training L2 learners to focus attention on the relevant acoustic cues signaling the differences between L2 sounds is an important factor promoting the successful L2 speech learning. Very few studies have investigated the specific contribution of attention-shifting ability to the L2 learners' phonological competence. To the best of our knowledge, only a few studies have explored the individual differences in attention shifting and L2 speech perception and production (Darcy et al., 2014, 2011; Mora & Safronova, submitted; Safronova & Mora, 2012b, 2013).

For example, in order to investigate attention control in L2 speech perception Safronova and Mora (2013, 2012b) and Mora and Safronova (submitted) developed a speech-based adaptation of Segalowitz and Frenkiel-Fishman's (2005) attention-shifting task. In this novel task 58 Catalan-Spanish participants were asked to shift focus of attention from one speech-based dimension (segmental duration: long vs. short) to another (voice quality: female vs. male) in the perception of vowel sounds. Shift costs were found to be unrelated to the perception of English /i/-/ɪ/ vowel contrast. The participants' lower error rates were significantly correlated with more accurate L2 vowel discrimination (accounting for 33.6% of unique variance). The results suggested that the ability to accurately foreground relevant acoustic information rather than the speed of attention shifting provided L2 learners with an advantage in L2 speech learning (Safronova & Mora, 2013).

In another recent study, Darcy et al. (2014) examined the role of attention control in 16 L1-Spanish learners of English and 18 L1-English learners of Spanish perception of L2 vowel and consonant contrasts measured via an ABX task. The participants' attention-shifting ability was measured using a speech-based attention-shifting task which required participants to perform two task sets alternating predictably every two trials. In the first subtask participants had to decide if nonwords started with a nasal consonant or not, whereas in the second subtask they had to judge if they were English sounding or not. The results showed that L1-English learners' lower shift costs were significantly related to more accurate perception of L2 vowel and consonant contrasts. However, they found no relationship between attention control and perception scores for L1-Spanish learners of English. The authors suggested that a more efficient attention control might facilitate the processing of relevant acoustic information and may lead to the development of accurate L2 phonetic categories. In a more recent study Darcy et al. (2015) found that attention control also operationalized as attention shift cost was unrelated to adult Korean learners' perception (N = 30) of English vowel /i/-/ɪ/, /u/-/ʊ/ and /æ/-/ɛ/, and consonant /p/-/f/ and /ɹ/-/l/ contrasts.

To summarize, attention control, and its attention-shifting function, in particular, can potentially play an important role in L2 learners' acquisition of perceptual phonological competence. However, findings so far do not yield a clear picture of the contribution of attention control in L2 speech learning. Specifically, it is still not well understood how accuracy and speed in attention shifting promote L2 learners' acquisition of L2 phonological contrasts. This is one of the aims of this dissertation.

3.5 Summary

This chapter has addressed the issue of the role of cognitive ability in L2 learning. Short-term memory and attention control are considered crucial cognitive skills underlying the processing of L2 input, and as a result the acquisition of new L2 knowledge. The L2 studies reviewed suggest that inter-learner variation in L2 learning can be attributed to learners' individual differences in phonological short-term memory and attention control.

Research on speech perception and memory suggests the existence of two distinct short-term memory storage systems responsible for maintaining phonological (i.e. categorical information) and acoustic (i.e. continuous information) properties of speech in working memory. Although the exact nature of acoustic short-term memory is still not clear, it may constitute a part of working memory and play an important role in the perception of L2 sounds.

Despite the growing evidence supporting the relationship between phonological short-term memory, acoustic short-term memory, attention control and L2 speech learning, little research has been done to investigate the role of these cognitive skills in L2 speech learning. As discussed in Chapter 2, L2 speech learning is a challenging task which requires L2 learners to deal with the influence of their L1 phonological system. In order to acquire L2 perceptual phonological competence L2 learners must learn to perceive acoustic-phonetic properties distinguishing L1 and L2 sounds, and contrasting L2 sounds (Best & Tyler, 2007; Flege, 1995, 2007; Flege et al., 2003).

This dissertation aims to add to previous L2 psycholinguistic research on L2 speech learning by investigating the role of phonological short-term memory, acoustic short-term memory and attention control in learners' cross-language speech perception and their acquisition of L2 perceptual phonological competence. On the basis of previous research outcomes we predict that greater phonological short-term memory, acoustic short-term memory and more efficient (accurate and fast) attention control are associated with learners' ability to perceive phonetic distance between L1 and L2 sounds and to acquire L2 phonological contrasts.

In order to investigate the role of L2 learners' cognitive ability and their perceptual phonological competence, we set out an experiment in which we related learners' phonological short-term memory, acoustic short-term memory and attention control abilities to their perception of a cross-language phonetic distance between L2 and L1 vowel sounds and the ability to discriminate between contrasting L2 vowel sounds. The detailed description of the methodology and the results obtained are reported in the following chapters.

Chapter 4

The study

This chapter presents the methodology of our study investigating the relationship between phonological short-term memory, acoustic short-term memory, attention control and L2 learners' perceptual phonological competence. Our aim was to relate the scores of the phonological short-term memory, acoustic short-term memory and attention control tasks to the scores of the L2 vowel perception tasks and estimate the contribution of these cognitive abilities to the variance in L2 perceptual phonological competence. As discussed in Chapter 1, previous research has identified factors such as age of onset of L2 learning, experience with the L2 and amount of L2 use, length of residence in an L2-speaking environment and L2 proficiency to be the strongest predictors of successful L2 speech learning. In the present study we also obtained measures for these L2 learning-related factors in order to facilitate analyses of individual differences in L2 perceptual phonological competence. This chapter aims at describing the design of the study by providing information about the participants' background as well as a detailed description of the tasks used to assess the participants' cognitive skills and their perception of L2 sounds. The chapter justifies the measures employed and describes the procedures followed to administer the experiment.

We address the issue of inter-learner variation in the acquisition of L2 sounds by testing Catalan-Spanish bilingual learners' perception of American English vowels, comprising /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ vowel contrasts, which are phonological in English and do not exist in Catalan and Spanish. Our choice of these English vowel contrasts is explained in Section 4.2.2 below.

4.1 Participants

4.1.1 L2 learners

The L2 participants in this study were 58 Catalan-dominant Catalan-Spanish bilingual EFL learners, who were undergraduate and graduate students at the University of Barcelona. They were selected from a larger pool of Catalan-Spanish bilinguals on the basis of two online questionnaires (see Appendices A and B and Section 4.2.1 for details). The participants were informed about the main purposes of the research, procedures as well as the potential benefits. They agreed to participate in the study voluntarily by signing a consent form. The undergraduate students were given course credit and graduate students were given a USB memory drive for their participation.

Table 2
Demographic characteristics of the group of L2 learners.

Measure	*N*	*M*	*SD*	*Lowest*	*Highest*
Age at testing (years)	45	20.2	3.0	18	31
AOL (years)[a]	45	6	2.7	2	15
Current daily use of Catalan (%)	45	60.9	15.7	30	90
Current daily use of Spanish (%)	45	21.6	13.5	0	45
Dominance in Catalan (BLP score)	45	62.9	36.9	11.2	150.1
Current daily use of English (%)	45	17.5	9.4	1	40
Current weekly use of English (hours)[b]	45	22.6	14	3	57
Average LoS (weeks)[c]	45	3.3	7.6	0	40
AmE exposure (%)	45	47.8	26.2	0	100
Self-estimated proficiency	45	6.8	1.1	3.50	8.75

[a] AOL = age of onset of L2 learning.
[b] L2 use measure included an amount of hours a week participants used English at home and with native English-speaking friends, watched films, listened to music and read books in English.
[c] LoS = length of stay abroad.

The participants were asked to perform two L2 speech perception tasks assessing the ability to perceive a cross-language phonetic distance between L2 and L1 vowels, and the L2 vowel discrimination skills. They also performed a battery of speech-based cognitive tasks measuring their phonological short-term memory, acoustic short-term memory and attention control abilities.

Out of 58 participants, only those who reported no speech/hearing disorders and obtained valid data in all the tasks were selected for further data analyses. The participants whose performance on the control conditions of the L2 vowel perception and cognitive tasks was below or above 2.5 standard deviations (SD) from the group mean were excluded from further analyses. Data from 45 participants (male = 9, female = 36, $M_{\text{age}} = 20.2$ years, $range = 18 - 31$ years, $SD = 3.0$) were considered valid in the analyses (see Table 2).

4.1.2 Native speakers

A group of 16 native speakers of American English were recruited to provide baseline data for the L2 vowel discrimination task. They were undergraduate and graduate students in Indiana University (Bloomington, USA). The study was advertised by the researcher in the students' classrooms. The students were informed about the purpose, experimental procedures, risks and benefits of the study.

The volunteers gave their oral consent and were asked to fill in an anonymous linguistic background questionnaire that was used to control for the participants' language background, such as American English-speaking parents, place of birth and residence, exposure to and use of other languages (see Appendix C). The participants were paid for taking part in the experiment.

The data from the participants who reported being bilinguals in another language ($N = 1$) or who performed on the control condition beyond 2.5 SD from the group mean ($N = 1$) were discarded. The data from 14 native speakers (male= 6, female=

8, $M_{\text{age}} = 24.2, range = 18 - 33$) were further analyzed and compared to the L2 learners' performance[1].

4.2 Instruments

4.2.1 Questionnaires

4.2.1.1 Bilingual language profile questionnaire

The participants in the group of L2 learners were selected by means of a modified version of the online Bilingual language profile (BLP) questionnaire (Birdsong, Gertken, & Amengual, 2012). The questionnaire was used as a tool for identifying participants' dominance in Catalan. The original BLP questionnaire was adapted to a specific bilingual environment in Catalonia and to the needs of the present study (see Appendix A)[2]. The questionnaire was administered in Catalan and consisted of five modules: *Personal information*, *Linguistic background*, *Language use*, *Linguistic competence* and *Attitudes* which aimed to elicit information about the participants' age of onset of acquisition of Catalan and Spanish, linguistic competence in each language, frequency of use and exposure and their personal attitudes towards Catalan and Spanish.

The global dominance score for each language was calculated by adding up the scores in four modules (*Personal information* module was not scored). The maximum score was 268 for each language (150 + 50 + 40 + 28 = 268). In order to ensure that each module received equal weighting of 67 points (268/4= 67) in the global dominance score, the total score in the modules was multipled by factors of 0.446 (*Linguistic background*), 1.34 (*Language use*), 1.675 (*Linguistic competence*) and 2.393 (*Attitudes*)[3].

[1]The baseline data were collected with the permission of Indiana University Institutional Review Board (IRB study 1401229361).

[2]The questionnaire was designed using Google Forms.

[3]The description of the scoring method can be found at https://sites.la.utexas.edu/bilingual/scoring-and-interpreting-the-results/

In order to obtain the participants' dominance index, the global score for Spanish was subtracted from the global score for Catalan. A BLP score near zero indicated balanced bilingualism whereas a more positive or a more negative score reflected the participants' dominance in Catalan or Spanish, respectively. Those participants who obtained a positive score (>0) in the BLP were asked to participate in the present study. The descriptives of the BLP scores are presented in Table 2.

4.2.1.2 Linguistic background questionnaires

The L2 learners selected on the basis of the BLP were further asked to fill in an anonymous online linguistic background questionnaire (see Appendix B). The questionnaire aimed at acquiring information about L2 learners' age of onset of learning English, amount of exposure and use of English, the variety of spoken English (British or American English), amount of exposure to British and American English and length of stay in an English-speaking country. These data were collected in order to explore the relationship between the above mentioned L2 learning-related factors and individual differences in L2 perceptual phonological competence (see Table 2 for descriptives).

The participants reported having started to learn English at about the age of 6 years ($SD = 2.7, range = 2-15$ years) and being exposed to English, primarily, through formal classroom instruction in a Foreign Language context, extracurricular classes (e.g., private language schools) and mass media.

Regarding L2 exposure and L2 use, the questionnaire asked the participants to estimate the amount of hours a week they used English at home, with native English-speaking friends, watching films, listening to music and reading books in English. The sum of self-estimated hours of weekly use of English was used as a measure of participants' L2 exposure in the present study. As shown in Table 2 participants reported using English in the above mentioned domains on average 22.6 hours a week ($SD = 14.0$).

The participants were also asked to report the variety of English they used and were more exposed to. Twenty-four percent of participants reported being exposed to British

English and American English 50% of their time, 40% and 36% reported being exposed more frequently to British English or American English, respectively. As shown in Table 2, participants reported being exposed to American English 47.8% of the time ($SD = 26.2$, $range = 0 - 100$). Twenty-six participants (58%) claimed they spoke English with a British accent and 19 participants (42%) believed they spoke English with an American accent. Since the present study focuses on the perception of American English, the possibility of an advantage of learners more often exposed to American English was taken into account by exploring the relationship between amount of exposure to American English and L2 perception measures.

Eighteen participants (40%) reported having lived in an English-speaking country for periods longer than two weeks. For most of those participants the stay abroad lasted from 2 weeks to one month ($N = 14$). Thirty-three participants (60%) reported they had never stayed abroad for more than 2 weeks. On average participants had lived in an English speaking country for 3.3 weeks ($SD = 7.6$). We assessed the effect of stay abroad on L2 learners' perceptual phonological competence.

The linguistic background questionnaire also asked participants to self-estimate their L2 proficiency in speaking, listening, reading, writing and pronunciation on a nine-point scale (1= *Very poor*, 9= *Near native*) and provide information about English proficiency level certificates (PET, FCE, CAE and CPE) obtained. The participants' average self-reported proficiency ranged from 3.50 to 8.75 ($M = 6.81$, $SD = 1.11$). Eleven participants (24%) reported holding PET or FCE certificates, nine participants (20%) having CAE or CPE certificates and twenty-five (56%) holding no certificate in English. Along with the scores of the vocabulary size tests (see Sections 4.2.6 and 5.3) these data suggest that the participants' proficiency level ranged from low-intermediate to advanced.

4.2.2 L2 perception tasks

In this study we used a perceptual assimilation task and a categorical vowel discrimination task to assess the participants' L2 vowel perception. The perceptual assimilation task served as an instrument to measure the participants' degree of perceived phonetic distance between L2 and L1 vowels, and the distance between perceptually similar L2 vowels in L2 learners' phonological system. We used this task to assess L2 learners' ability to establish new phonetic categories for L2 sounds, which, as discussed earlier in the dissertation (see Chapter 2), determines learners' success in L2 speech learning. The categorical ABX vowel discrimination task was used as a measure of the participants' categorical perception of L2 vowel contrasts, that is, their L2 perceptual phonological competence.

The L2 perception tasks targeted the participants' perception of five General American English vowel contrasts /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/. As discussed in Chapter 2, L1-Catalan learners of English may have difficulties in acquiring these English vowel contrasts due to the differences in English and Catalan vowel inventories and the perceived phonetic similarity between Catalan and English vowels. The present section provides a detailed description of the L2 vowel perception tasks.

4.2.2.1 Perceptual assimilation task

The participants' cross-language vowel perception was assessed by means of a perceptual assimilation task (Bungaard-Nielsen et al., 2011; Cebrian, 2009; Cebrian et al., 2011; Guion et al., 2000; Lengeris, 2009). In this task participants had to identify L2 sounds in terms of L1 categories and provide goodness-of-fit ratings. As discussed in Chapter 2, the perceptual assimilation task has been widely used to assess the degree of perceived phonetic similarity between L2 and L1 sounds (Bohn, 2002; Lengeris, 2009; Strange, 2007). The task has also been used to predict the discriminability of contrasting L2 sounds by providing an estimate of how often two members of an L2 sound contrast are assimilated to a single L1 phonetic category (Levy, 2009b). Perceptual assimilation scores may indicate the ability to establish distinct phonetic categories for L2 sounds.

In the present study the perceptual assimilation task required participants to identify six American English vowels /i/, /ɪ/, /ɛ/, /ɑ/, /ʌ/ and /æ/ as instances of Catalan vowel categories and rate English vowels as regards their goodness of fit to the Catalan vowel categories. As an experimental control, four Catalan vowels /i/, /e/, /ɛ/ and /a/ were included for identification and rating. The English and Catalan vowels were embedded in /bVs/ consonantal contexts which created CVC monosyllabic stimuli for both languages (Cebrian, 2009). The English and Catalan stimuli were *beace* (/i/), *biss* (/ɪ/), *bess* (/ɛ/), *boss* (/ɑ/), *buss* (/ʌ/), *bass* (/æ/), and *biss* (/i/), *bèss* (/ɛ/), *béss* (/e/) and *bass* (/a/), respectively. The stimuli were examined by two phonetically trained native speakers (i.e. a native speaker of English and a native speaker of Catalan) to ensure that the stimuli conformed to the phonotactics and syllabification rules of both languages.

The English stimuli were elicited in a carrier phrase *It rhymes with ____, I say ____. I say ____ again* (e.g., *It rhymes with kiss, I say biss. I say bis again*). The stimuli were recorded by three volunteer male native speakers of American English ($M_{\text{age}} = 30$ years, $range = 28 - 33$ years). The speakers filled in an anonymous linguistic background questionnaire eliciting information about their place of birth and residence, proficiency in other languages and parents' mother tongue (see Appendix C). The speakers came from several places in the USA (Vermont, Oklahoma and Connecticut) who at the moment of the recording were living in Bloomington (Indiana, USA). Each speaker produced eight repetitions of the phrases at a normal speaking rate and with falling intonation. The recordings (Praat and Edirol UA-25 USB Audio Capture device) were made in a soundproof booth in the L2 Psycholinguistics Lab at Indiana University (Bloomington, Indiana, USA).

The Catalan vowel stimuli were elicited from three male native speakers of Eastern Catalan living in the Barcelona area [4]. The Catalan speakers ($M_{\text{age}} = 43$ years, $range = 35 - 52$ years) were selected on the basis of the language background questionnaire, where they reported speaking Catalan as their first and dominant language and using Catalan 75-100% on a daily basis (see Appendix D). Each Catalan speaker pro-

[4] Two speakers were from Mataró and one speaker was from Badalona.

duced eight repetitions of the stimuli embedded in a carrier phrase *Rima amb ____, ara dic ____. Ara dic ____*, meaning "It rhymes with ____, I say ____. I say ____ again" (e.g., *Rima amb pis, ara dic biss. Ara dic biss*). The Catalan stimuli were digitally recorded using Marantz PMD660 in sound-attenuated conditions.

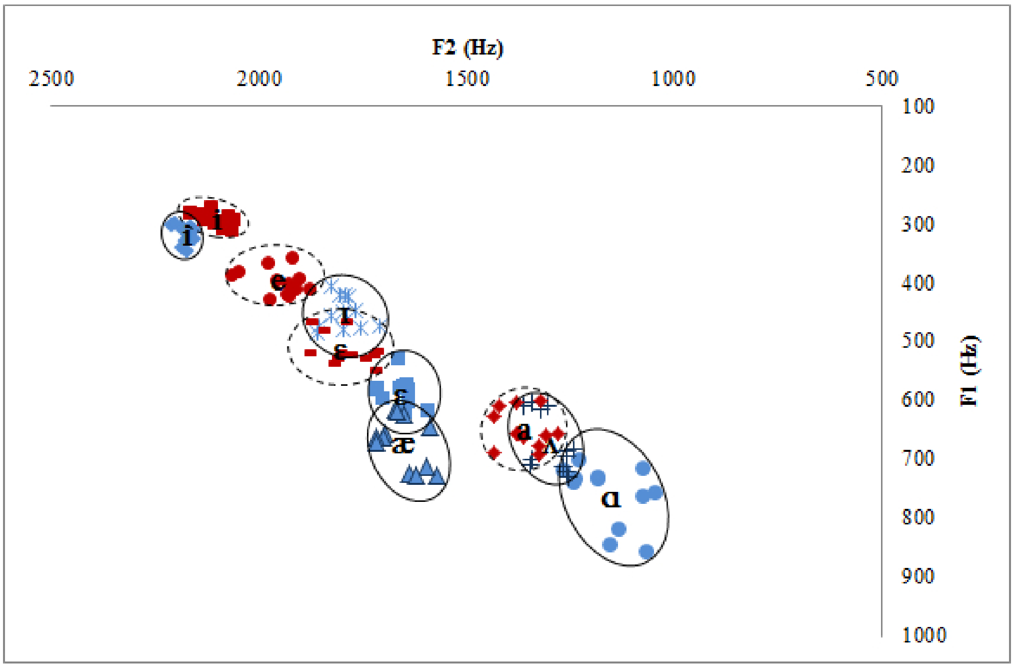


Figure 6. F1 and F2 frequencies (in Hz) measured within a 20 ms window positioned at the vowel peak intensity in each of the English (blue) and Catalan (red) vowel stimulus for the perceptual assimilation task. Phonetic labels are located at the mean F1 vs. F2 values. The ellipses surrounding English (solid line) and Catalan (dashed line) are drawn for illustration purposes only.

All recordings were digitized at a 44.1 kHz sampling rate with 16-bit resolution. The CVC stimuli were extracted from the carrier sentences and were normalized for amplitude (70 dB). The best four tokens per speaker and language were selected on the basis of auditory judgment and acoustic measurements (see Appendix E, Table E.1). In order to neutralize potential effects of cross-language differences in consonants on the participants' perceptual judgment of CVC nonwords, the saliency of nonword final /s/ in both languages was smoothed (smoothing factor of 7 points). The Catalan stimuli were further processed by removing the prevoicing of the initial /b/ consonant to minimize the differences in prevoicing of this consonant in the two languages[5]. Figure 6 plots the F1 and F2 frequences of the English and Catalan vowels used in the perceptual assimilation task. The mean F1 and F2 frequencies as well as mean vowel duration of the English and Catalan vowel stimuli as produced by three male native speakers in each language are presented in Table 3[6].

Table 3

Mean F1 vs. F2 frequencies (in Hz) and mean duration (in ms) of the vowel stimuli in the perceptual assimilation task.

	Vowel	**F1**	**SD**	**F2**	**SD**	**Duration**	**SD**
English	/i/	316	14.8	2176	17.1	119	5.2
	/ɪ/	450	27.9	1797	42.7	107	6.2
	/ɛ/	589	25.9	1659	35.7	109	4.7
	/æ/	670	42.9	1653	50.9	135	5.0
	/ʌ/	672	48.8	1296	43.9	111	5.8
	/ɑ/	761	51.8	1154	77.5	138	3.5
Catalan	/i/	291	11.4	2103	32.9	132	9.8
	/e/	395	22.5	1954	57.1	131	5.9
	/ɛ/	511	26.6	1802	57.8	132	6.7
	/a/	649	31.9	1362	51.9	131	6.5

A total of 120 stimuli, which included four tokens for each of the six English and the four Catalan vowels as produced by three different male speakers in each language, were randomly presented to the participants using OpenSesame experiment builder and presentation software (Mathôt, Schreij, & Theeuwes, 2012). The stimuli were presented to the participants for two kinds of auditory judgment. They were first asked to identify each vowel stimuli as one out of seven Catalan vowel categories. A set of seven line drawing pictures (see Figure 7) used as representative exemplars for seven Catalan vowels were presented to the participants as all possible response alternatives: *serp* (/e/) "snake", *gos* (/o/) "dog", *nas* (/a/) "nose", *llit* (/i/) "bed", *foc* (/ɔ/) "fire", *set* (/ɛ/) "seven", and *suc* (/u/) "juice". The pictures were used to represent the response categories instead of

[5] In contrast with English voiced stops Catalan are prevoiced because voicing begins during the stop closure, before the stop release, which results in negative voice onset time (Ladefoged, 2012; Recasens, 1996).

[6] F3 frequency related to lip rounding is not presented because it is non-distinctive in English and Catalan.

orthographic labels (e.g., Guion et al., 200; Levy, 2009), in order to avoid the possible effect of the orthographic similarity of labels on the participants' judgment of similarity between speech sounds (Bohn, 2002).

After having selected the Catalan category best matching the L2 vowel stimulus, participants were asked to listen to the token again by clicking on the button *Escolta* ("Listen") displayed on the screen, and to rate the token for a goodness of fit to the selected Catalan category using a seven-point scale scale (1= *totally different* to 7= *identical*). Next stimulus was presented by participant's clicking on *Següent* ("Next") button. If necessary participants could take a short break after the first 60 trials.

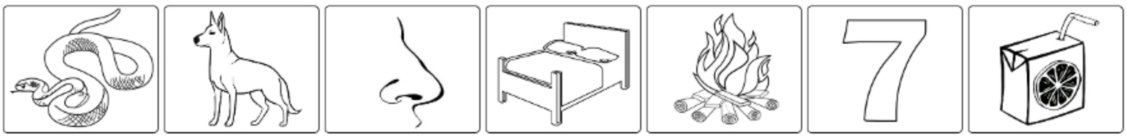


Figure 7. L1 response categories in the perceptual assimilation task.

Participants were instructed in Catalan, both orally by the researcher and visually by the instructions displayed on the screen (see Appendix E, Figure E.1). The instructions explained the task procedure, the use of the response categories and the rating scale. Participants were asked to listen to the stimuli over the headphones and provide their responses with a mouse click by ticking the boxes displayed on the screen. Prior to the test block a six-trial practice block was given to familiarize participants with the task. Participants' responses were analyzed in terms of the percentages of identification of English vowels as instances of Catalan categories as well as their goodness of fit ratings.

The main purpose of the task was to assess participants' perception of a cross-language phonetic distance between English and Catalan vowels, and the degree of perceptual overlap of English vowels in the /i/- /ɪ/, /ɪ/- /ɛ/, /ɑ/- /ʌ/, /ɑ/- /æ/ and /ʌ/- /æ/ vowel contrasts. The measure of the participants' degree of perceived phonetic distance between L2 and L1 vowels was a perceived cross-language phonetic similarity score (henceforth *perceived phonetic similarity score*). This score was obtained by computing *fit indexes* for L2 vowels identified as instances of L1 vowel categories (Guion et al., 2000). First, the fit indexes were calculated for each target L2 vowel by multiplying its highest proportion of identification by the corresponding goodness-of-fit rating. For example, English /i/ and /ɑ/ were identified as Catalan /i/ and /a/, respectively, 100% of the time. English /ɪ/ was identified at 50% and 50% rates as Catalan /i/ and /e/. English /ɛ/ was assimilated to Catalan /ɛ/ and /a/ in 50% and 42% of the cases, respectively. English /ʌ/ was identified as Catalan /a/ and /o/ in 70% and 30% of instances, respectively. English /æ/ was assimilated to Catalan /a/ at 75% and to Catalan /ɛ/ at 25% rates. In a given example the highest percentages of identification for English /i/, /ɑ/, /ɪ/, /ɛ/, /ʌ/ and /æ/ are 100, 100, 50, 50, 70 and 75, respectively.

The proportions of identifications were then multiplied by the goodness of fit rating. If a L2 vowel was identified as two L1 Catalan categories at similar rates like English /ɪ/ in the example, the fit index for that L2 vowel was calculated on the basis of the highest goodness-of-fit rating. The total perceived phonetic similarity score was computed by averaging fit indexes of six target L2 vowel categories and converting it to a percentage score. In the example described above the perceived phonetic similarity score would be calculated as follows: CLPS = mean(1x6.5, 0.50x5, 1x7, 0.50x4.5, 0.70x6, 0.75x3.5)x100/7 = 59.7%. It was predicted that lower perceived phonetic similarity scores would indicate L2 learners' better ability to perceive phonetic distance between L2 and L1 vowels and, thus, reflect L2 learners' better ability to establish new phonetic categories for L2 sounds.

In order to measure the distance between perceptually similar L2 vowels in the L2 learners' phonological system, a cross-language assimilation overlap score (henceforth *assimilation overlap score*) was obtained. The score was computed using the *cross-language assimilation overlap* method which assesses the degree of perceptual overlap between two L2 vowels in a contrast assimilated to a single L1 vowel category (Levy, 2009). The method introduced by Levy (2009) is based on the estimation of how often two contrasting L2 vowels are identified as instances of the same native category. The assimilation overlap score was calculated for each participant and each of the five English vowel contrasts which participants were later asked to discriminate in the L2 vowel discrimination task (see Section 4.2.2.2). It was computed by adding up the smaller percentages of responses when two members of a vowel contrast were perceptually assimilated to a particular Catalan vowel. For example, for the /i/ - /ɪ/ vowel contrast, the participant ID25 assimilated English /i/ to Catalan /i/ at 92% rate and assimilated English /ɪ/ to Catalan /i/ 8% of the time. Therefore, the perception of English /i/ and /ɪ/ for this participant overlapped in 8% of the cases (i.e. the smaller percentage). In addition, both /i/ and /ɪ/ were perceived as closest to Catalan /e/ for an overlap of 8%. Then all the overlap percentages for a given vowel contrast were added up. Thus, the overlap score for the English /i/ - /ɪ/ contrast was 16%.

A single assimilation overlap score was calculated by averaging the total overlap score for the five English vowel contrasts. Lower scores were predicted to reflect lower degree of assimilation of L2 vowels to a single L1 category, thus, to indicate a better ability to establish new phonetic categories for L2 vowels.

4.2.2.2 ABX vowel discrimination task

The participants' accuracy in the perception of L2 phonological contrasts was assessed by means of a categorical ABX discrimination task. As discussed in Chapter 2, this task has proved to be a reliable method to measure listeners' between- and within-category discrimination of speech sounds (Fujisaki & Kawashima, 1970; Pisoni, 1973, 1975). The task has been often used in previous L2 speech perception studies to assess L2 learners' phonological competence, that is, the extent to which L2 phonetic categories have been established (Darcy et al., 2014, 2015; Gottfried, 1984). In an ABX task participants hear three stimuli in a row and decide if the third stimulus (X) is the same as the first one (A) or the same as the second one (B).

The task was designed to asses participants' perception of five test English vowel contrasts (/i/- /ɪ/, /ɪ/- /ɛ/, /ɑ/- /ʌ/, /ɑ/- /æ/ and /ʌ/- /æ/). A control vowel contrast /i/- /ɑ/ was included to ensure that the participants did the task properly. The test and control vowel contrasts were embedded into six consonantal contexts with a ˈCVCVC structure (see Table 4). The nonword initial consonants and consonants following the test vowels varied in their place and manner of articulation. The consonants following the test vowels were all voiceless in order to avoid natural lengthening of vowels preceding voiced consonants in English. The stimuli were examined by a native speaker of American English to make sure that they conformed to English phonology.

Table 4
Stimuli for the ABX vowel discrimination task.

Vowel contrast	Condition	Consonantal contexts	Example stimuli	
			Item A	**Item B**
/i/ vs. /ɪ/	Test	[ˈp_kəs]	[ˈpɑkəs]	[ˈpʌkəs]
/ɪ/ vs. /ɛ/	Test	[ˈk_fɪv]	[ˈkifɪv]	[ˈkɪfɪv]
/ɑ/ vs. /ʌ/	Test	[ˈs_tʃən]	[ˈsitʃən]	[ˈsɑtʃən]
/ɑ/ vs. /æ/	Test	[ˈt_sɪʃ]	[ˈtæsɪʃ]	[ˈtʌsɪʃ]
/ʌ/ vs. /æ/	Test	[ˈl_tɪf]	[ˈlɑtɪf]	[ˈlætɪf]
/i/ vs. /ɑ/	Control	[ˈm_kət]	[ˈmɪkət]	[ˈmɛkət]

The ˈCVCVC nonwords were embedded in a carrier phrase that included a real English word to illustrate how the test vowels should be pronounced (e.g., *It sounds like puppet. I say puckus. I say puckus again.*). Three male native speakers of standard American English living in Bloomington (Indiana, USA) volunteered in the recording session ($M_{\text{age}} = 32$ years). The speakers filled in a linguistic background questionnaire (see Appendix C) eliciting information about their place of birth, parents' mother tongue and linguistic background. Then they recorded the stimuli at a normal speaking rate on a falling intonation. The stimuli were digitally recorded (Praat and Edirol UA-25 USB Audio Caprture device) in a soundproof booth at a sampling rate of 44.1 kHz with a 16-bit resolution on a mono channel in the L2 Psycholinguistics Lab at Indiana University (Bloomington). The target stimuli were extracted from the recorded carrier sentences and were normalized for amplitude (70 dB) using Praat speech analysis software (Boersma & Weenink, 2012).

The best tokens per speaker selected on the basis of auditory judgments and acoustic measurements were included as stimuli in the task. Figure 8 plots the average frequencies of the six test vowels. The description of the mean vowel stimuli duration per consonantal context is presented in Table 5. For more detailed information about the stimuli such as F0, F1, F2 average values and vowel duration for each speaker see Appendix F, Table F.1.

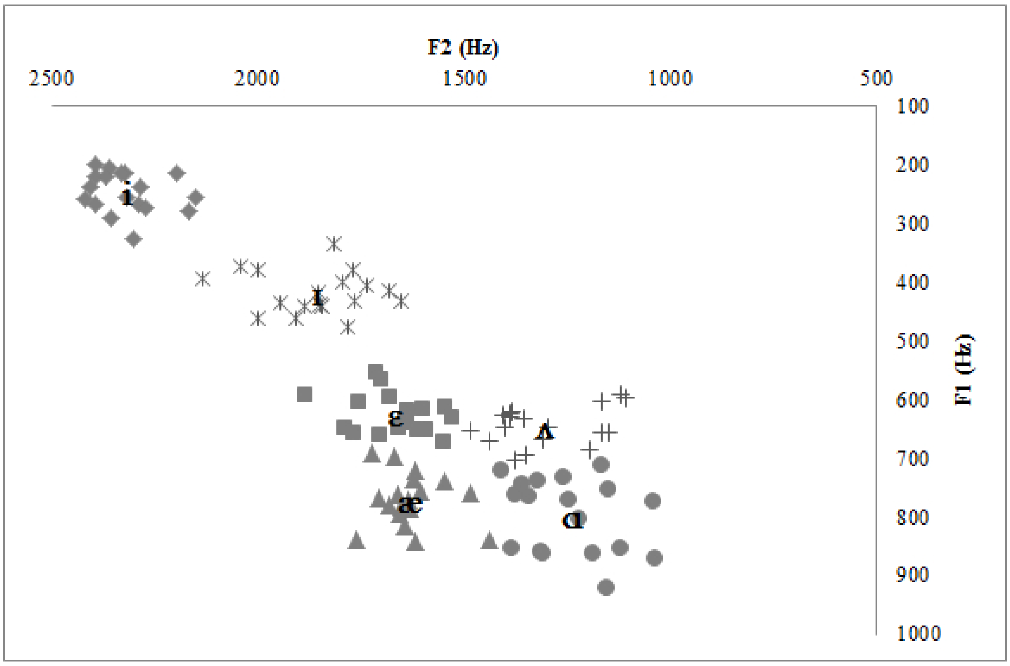


Figure 8. F1 vs. F2 formant frequencies for six tested vowels produced by three male speakers of American English.

The stimuli were then organized into experimental trials where each stimulus A, B and X, presented with an interstimulus interval of 1200 ms, was spoken by a different male voice. Such trial design was chosen in order to force participants to rely on abstract phonological representations rather than on auditorily detectable differences between the speech stimuli (Flege, 2003; Guion et al., 2000; Werker & Logan, 1985). To counterbalance order effects each vowel contrast was presented in four possible permutations: ABA, ABB, BAA and BAB, yielding a total of 144 trials (6 vowel contrasts x 6 consonantal contexts x 4 ABX combinations).

Participants were told that in this task, testing their perception of English vowels, they would hear three English nonwords where the first nonword (A) and the second nonword (B) were always different and that the task was to decide if the third nonword (X) was the same as the first or the second nonword. They were instructed to respond as accurately

Table 5
Vowel duration (in ms) of the stimuli in the ABX vowel discrimination task (SD in parentheses).

Vowel	**Consonantal context**						**Mean vowel duration (ms)**
	[ˈp_kəs]	[ˈk_fɪv]	[ˈs_tʃən]	[ˈt_sɪʃ]	[ˈl_tɪf]	[ˈm_kət]	
/i/	96 (1.0)	86 (6.0)	96 (4.0)	75 (16.0)	93 (3.1)	94 (1.0)	90 (4.1)
/ɪ/	76 (2.3)	65 (13.3)	61 (4.0)	58 (3.1)	63 (3.8)	74 (8.0)	66 (1.2)
/ɛ/	71 (3.8)	68 (3.8)	77 (4.2)	65 (1.0)	78 (6.6)	87 (8.3)	74 (2.3)
/æ/	100 (13.4)	111 (18.0)	118 (18.2)	102 (14.2)	118 (9.1)	125 (12.7)	112 (13.4)
/ʌ/	74 (3.5)	63 (3.8)	74 (6.4)	59 (4.0)	67 (6.1)	69 (1.0)	68 (0.7)
/ɑ/	87 (9.5)	99 (8.5)	128 (25.7)	103 (11.2)	111 (15.6)	111 (16.5)	106 (13.8)

and as fast as possible by pressing left or right Shift keys. The key assignment was reminded to participants by the picture displayed on the screen (see Appendix F, Figure F.1). Participants were given 3000 ms to respond. The next trial was presented 1500 ms after response or after 3000 ms if no response was given.

In order to familiarize participants with the procedure, the task started with a four-trial practice block in which visual feedback was provided. Feedback consisted of a picture of a tick or a cross appearing in the middle of the screen, or the string *Too slow!* if the 3000 ms time limit had elapsed. The experiment began after the participants had done the practice block and reported having understood the instructions. Trials in the test block were presented randomly in two sub-blocks of 72 trials with a short pause in between. In order to raise participants' interest and motivation at the end of the practice and test blocks they received visual feedback consisting of their average reaction time and accuracy score.

To avoid the potential effects of the differences in hardware performance on the reliability of the reaction time data, participants performed the task on the same PC. OpenSesame software recorded participants' responses and reaction times. The reaction times for the correct responses were screened for each participant and those above and below 2.5 SDs from the overall participant's mean reaction times were not considered for analysis (2% of total reaction times). The percentage of correct responses (henceforth *discrimination accuracy scores*) and average reaction times (henceforth *discrimination reaction time scores*) on test vowel stimuli were calculated and served as measures of the participants' L2 vowel discrimination ability. It was hypothesized that more accurate and effortless performance (i.e. higher accuracy scores and lower reaction time scores) would reflect the participants' higher level of competence in the perception of L2 phonological contrasts.

4.2.3 Phonological memory task

In the present study phonological short-term memory is understood as a temporary storage for phonologically encoded (i.e. categorized) verbal information of the speech stimuli. Participants' phonological short-term memory capacity was assessed using a Catalan version of the serial nonword recognition task created by Cerviño-Povedano and Mora (2011). In this task participants hear a pair of sequences of Catalan nonwords and have to decide if the order of the nonwords in the sequences is the same or different. The sequences of nonwords increase in length as the task progresses, thereby placing more demands on the test takers' phonological short-term memory capacity.

Some of the studies on phonological short-term memory capacity reviewed in Chapter 3 used the serial nonword recognition task, administered in participants' L1, as a reliable measure of phonological short-term memory capacity (O'Brien et al., 2006; O'Brien, Segalowitz, Freed, & Collentine, 2007; Cerviño-Povedano & Mora, 2011; Isaacs & Trofimovich, 2011). As Cerviño-Povedano and Mora (2011) explain, the serial nonword recognition task can be considered a better measure of the phonological short-term memory capacity than, for instance, nonword repetition or digit span, due to its avoiding articulatory constraints and minimizing lexical knowledge effects on the subjects' performance.

The stimuli in the phonological short-term memory task came from Cerviño-Povedano and Mora (2011). The task consisted of 24 pairs of 144 Catalan consonant-vowel-consonant (CVC) nonword sequences organized into three blocks of five-, six- and seven-item sequence pairs (see Appendix G, Table G.1). The nonwords in each sequence were separated by 300 ms of silence and presented in a sequence in such a way that all nonwords within a sequence contained a different vowel as well as a variety of consonants. Every block consisted of eight trials containing four same sequence pairs of two identical nonword (NW) sequences (e.g., NW1, NW2, NW3, NW4, NW5 - NW1, NW2, NW3, NW4, NW5) and four different sequence pairs in which two of the nonwords in a second sequence switched positions (e.g., NW1, **NW2**, **NW3**, NW4, NW5 - NW1, **NW3**, **NW2**, NW4, NW5). The position of the first and the last items in a sequence always remained constant to minimize the salience of the transposed items (Cerviño-Povedano & Mora, 2011: 57). The interstimulus interval between the two sequences in a sequence pair was 1000 ms. *Same* and *different* sequence pairs within each block were randomly presented to the participants but the blocks were presented according to the sequence length. In order to familiarize participants with the task a short four-trial practice block preceded the experimental blocks.

The task was administered through OpenSesame. The instructions were given to participants orally by the researcher and were also displayed on the screen (see Appendix G, Figure G.1). Participants were asked to decide as fast and as accurately as possible if the order of the Catalan nonwords in two sequences was the *Same* or *Different* by pressing left or right Shift keys, respectively. The task started once participants confirmed that they had understood the instructions and had completed the practice block. The measure of the participants' phonological short-term memory capacity was a weighted score, which takes into account the greater difficulty of correctly identifying pairs of longer sequences (O'Brien et al., 2007; Cerviño-Povedano & Mora, 2011). The measure was obtained by assigning five, six and seven points to the correct responses at five-, six- and seven-item sequences, respectively, for a total score of 144, which was then converted into a percentage score.

4.2.4 Acoustic memory tasks

In this study acoustic short-term memory is defined as an individual's storage for acoustic properties of speech stimuli at the acoustic (i.e. a pre-phonological) level. This ability was assessed through two versions of an auditory recognition task. In one version, participants were presented with two sequences of speech stimuli and had to decide if the sequences were the same or different (Tanaka & Nakamura, 2004). In the other version, participants were presented with a single sequence of speech stimuli and had to judge if a probe item was a member of the sequence or not (Prosser, 1995). The stimuli presented for recognition were nonwords that had been manipulated through frequency rotation so that they kept certain speech-like features (duration and pitch) but could not be encoded phonologically. Frequency rotation was implemented by inverting the speech spectrum around a center frequency (also known as rotated speech).

As discussed in Chapter 3, spectrally-rotated speech (see Figure 9) is as complex as normal speech (i.e. it preserves pitch, duration, intensity, voicing of the consonants and formant-like acoustic features of vowels) but lacks intelligibility and cannot be encoded phonologically (Blesser, 1972; Scott et al., 2000). These characteristics of rotated speech make it adequate for testing the storage capacity of acoustic memory (Safronova & Mora, 2012a). It was hoped that rotated speech would force participants to rely exclusively on their acoustic memory due to the unavailability of phonological coding and storage.

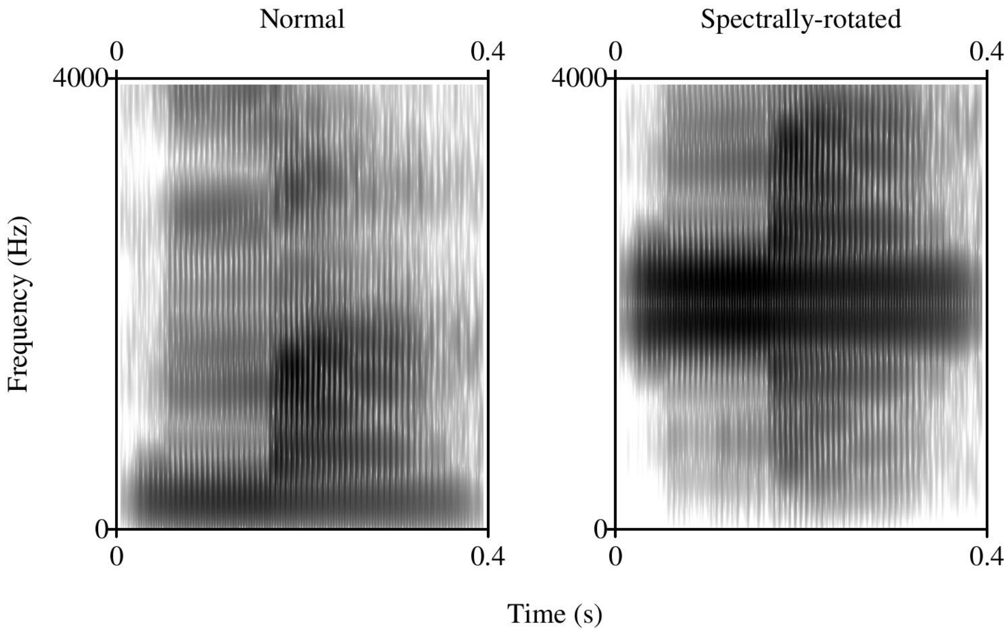


Figure 9. Spectrograms of normal Catalan /ma/ syllable (left) and its spectrally rotated version (right).

The two acoustic memory tasks were a serial sound recognition task and a target sound recognition task. The tasks aimed to test participants on their ability to hold the acoustic properties of speech in short-term memory. We assumed that the ability to hold a larger amount of acoustic information in memory would enhance a more accurate perception of L2 speech sounds.

4.2.4.1 Serial sound recognition task

The serial sound recognition task is based on a whole-display recognition paradigm (Rouder, Morey, Morey, & Cowan, 2011) and is similar in nature to the serial nonword recognition task used in the present study for assessing phonological short-term memory capacity. In the serial sound recognition task participants hear two sequences of speech-like sounds (with sequences increasing in length) and have to decide whether the two sequences are the same or different as regards the order of the sounds in both sequences (Mora & Safronova, submitted; Safronova, 2011; Safronova & Mora, 2012a).

The stimuli in this task were 96 Catalan CV syllables consisting of five Catalan vowels (/i/, /e/, /a/, /o/ and /u/) produced in a variety of consonantal contexts. The stimuli were selected from a larger list of 140 Catalan syllables (a combination of the seven stressed vowels of Catalan vowels and twenty consonants) spoken by a phonetically trained Catalan native speaker and digitally recorded (Marantz PMD660) in a sound attenuated booth at a 44.1 kHz sampling rate frequency with a 16-bit resolution. The recorded stimuli were pre-processed by normalizing amplitude (70 dB), and applying a linear amplitude ramp (10 ms) to both ends of the sound files (Owren, 2008). To avoid participants' reliance on differences in the CVs' duration, which may be a perceptually salient feature in this task, the duration of the CV syllables were shortened/lengthened to the mean duration of the stimuli in a trial using the PSOLA algorithm in Praat. The stimuli were further processed using a speech rotation technique (Blesser, 1972; Scott et al., 2000; Scott, Rosen, Beaman, Davis, & Wise, 2009). The technique involved low-pass filtering the original speech stimuli at 4000 Hz and applying spectral inversion at 2000

Hz, after which the modified speech stimuli were low-pass filtered again at 3800 Hz. This acoustic manipulation, performed using a rotation script for Praat (see Appendix H), produced unintelligible “alien-sounding” stimuli that preserved the acoustic complexity of normal speech.

The rotated stimuli were then organized into three-, four- and five-item sequence pairs with 200 ms interstimulus intervals between the items in a sequence and 750 ms between the two sequences in a pair (see Appendix H, Table H.1). To ensure the discriminability of the rotated items the CV syllables in a sequence were distributed on the basis of saliency of the consonants (manner, place of articulation and voicing) and vowel quality (frontness/backness). For example, each sequence consisted of CV syllables beginning with various consonants (stop, affricate, fricative, nasal, lateral and trill). None of the syllables in a sequence had the same consonant. The CV sequences did not comprise two or more adjacent alveolar, bilabial, dental or velar consonants, and two or more voiced or voiceless consonants.

Eight trials at each sequence length consisted of four *same* and four *different* sequence pairs. The latter were created by changing the position of one of the syllables in the second sequence (see Figure 10). The sequences were distributed in three test blocks of three-, four- and five-item sequences. Eight trials within each test block (for a total of 24 test trials) were randomly presented to participants with a 4000 ms delay upon response to give them time to recover from the previous trial.

Participants were informed that in this task they would hear a pair of sequences consisting of “strange” sounds and would have to decide whether the two sequences were the same or different by pressing the left or right Shift key, respectively. The key assignment was reminded to the participants by two pictures displayed on the screen (green and red rectangles with the labels *same* and *different* written on them). They were instructed to be as accurate as possible and to respond within five seconds (see Appendix H, Figure H.1).

A four-trial practice block was provided to familiarize participants with the stimuli and procedure. In the practice block participants received visual feedback that consisted

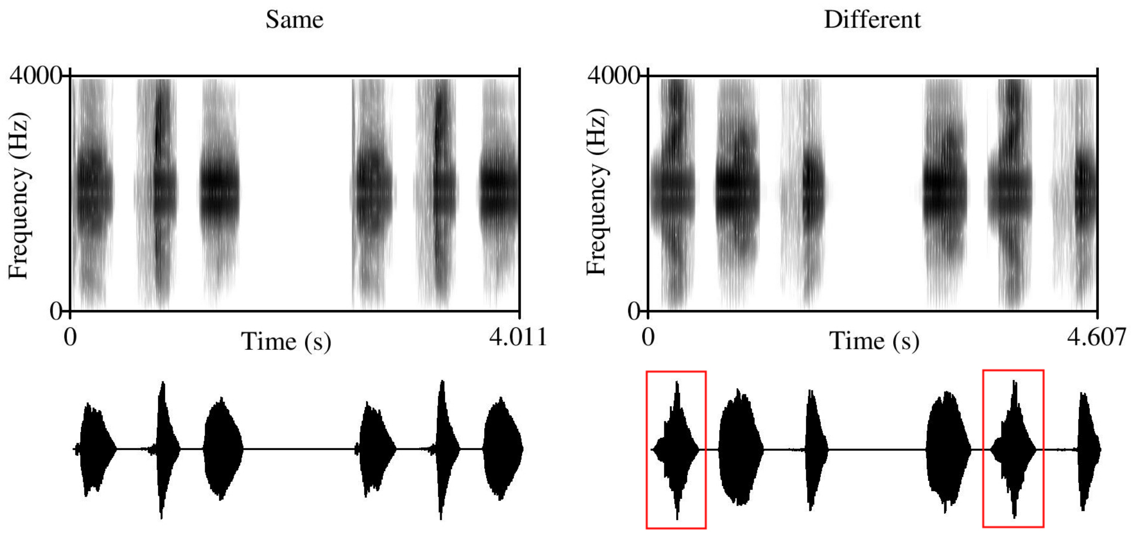


Figure 10. Sample trials of “same” (left) and “different” (right) three-item sequences in the serial sound recognition task.

of a picture of a tick (correct) or a cross (wrong) appearing in the middle of the screen. The task began after the participants had performed the practice block and reported having understood the instructions. At the end of the task visual feedback consisting of participants’ percentage of correct responses was displayed on the screen. Similar to the scoring method used in the phonological short-term memory task, the measure of the participants’ acoustic short-term memory capacity was a weighted score (out of 96) converted into a percentage score.

4.2.4.2 Target sound recognition task

The target sound recognition task was based on a single-probe recognition paradigm (Li et al., 2013; Prosser, 1995; Rouder et al., 2011). Participants had to decide if a sequence of sounds comprised a target sound presented after it (with sound sequences increasing in length as the task progressed). The task was assumed to tap on not only the ability to keep in memory sequences of acoustically complex items increasing in length, but also the ability to retrieve one of the items after a relatively long interstimulus interval of 3000 ms. Therefore, participants with an ability to store more acoustic information for longer periods of time were predicted to perform better on this task.

The stimuli in this task were 101 Catalan CV syllables. The stimuli recording, pre-processing and spectral rotation method were those described in Section 4.2.4.1. In order to neutralize the effect of the stimuli duration, the original CV syllables were shortened/lengthened to the mean duration of all CV stimuli (400 ms) in this task using PSOLA algorithm in Praat. This stimuli duration manipulation procedure was performed to ensure the equal duration of the target sound in each trial at each sequence length. The rotated syllables were separated by 200 ms interstimulus interval and were distributed in two-, tree- and four-item sequences (see Appendix H, Table H.2) followed by 3000 ms silence and a target sound which matched one or none of the items in the sequence (see Figure 11). Three test blocks containing eight trials (four *Yes* and four *No* trials at each sequence length) were created.

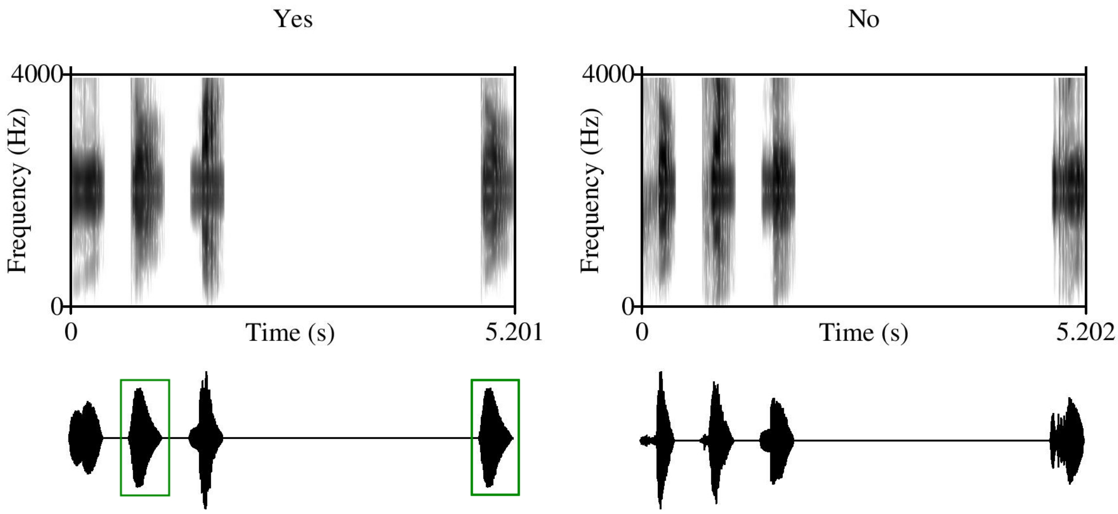


Figure 11. Sample trials of "yes" (left) and "no" (right) three-item sequences in the target sound recognition task.

In this task the trials in each test block (for total of 24 test trials) were randomly presented with 4000 ms delay upon response to give participants time to recover from the previous trial. In each trial the presentation of a sequence was cued by a picture of a music folder which signaled the participants to listen to the sequence and remember it. After a 1500 ms pause a picture of a loudspeaker, which served as a cue to the target sound to be played shortly, appeared on the screen and a target sound was presented 1500 ms after that (see Figure 12).

Participants were told that in this task they would hear a sequence of “strange” sounds followed by a target sound presented to them through headphones (see Appendix H, Figure H.2). They were instructed to follow the visual cues and judge whether or not the target sound belonged to the sequence by pressing the assigned key. The key assignment was reminded to the participants by pictures of green and red rectangles with the labels *Yes* and *No* written on them. Participants were asked to be as accurate as possible (or guess if necessary) and respond within five seconds. Before the task started a two-trial practice block had been provided to familiarize participants with the stimuli and procedure. In the practice block participants received visual feedback (picture of a tick or a cross standing for correct or incorrect answer, respectively). No feedback was provided during the test block. At the end of the task visual feedback consisting of participants’ percentage of correct responses was displayed on the screen.

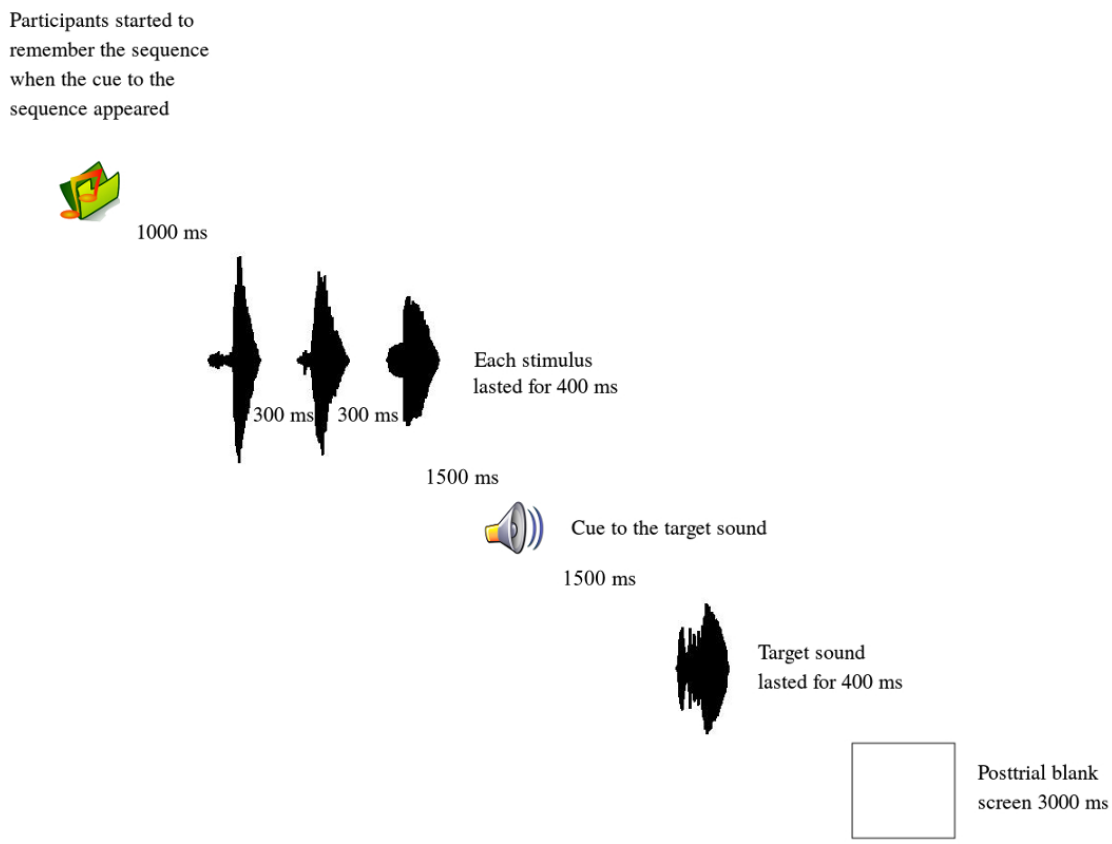


Figure 12. Illustration of the target sound recognition task procedure.

The scoring method used to obtain a measure of the acoustic short-term memory capacity was similar to the method used in the phonological memory task and acoustic memory serial sound recognition task. The weighted score was computed by assigning scores of two, three and four points to the correct responses on two-, tree- and four-item sequences, respectively. The total score (out of 72 points) was converted into a percentage score. It was predicted that higher scores would reflect a better ability to preserve a larger amount of acoustic information in acoustic short-term memory.

4.2.5 Attention-shifting task

The participants' phonological attention control was assessed through a speech-based attention-shifting task (Mora & Safronova, submitted; Safronova, 2011; Safronova & Mora, 2013). This task, based on the predictable alternating-runs task-switching paradigm (Rogers & Monsell, 1995; Segalowitz & Frenkiel-Fishman, 2005), measured the participants' ability to rapidly and accurately shift focus of attention between two acoustic speech-related dimensions - voice quality (female vs. male) and segmental duration (long vs. short). In order to complete this task, participants have to be able to accurately and rapidly bring to the perceptual foreground the speech dimension required by the current sub-task (e.g. sound duration) while inhibiting the irrelevant dimension (e.g. voice in which the sound was produced). We chose the voice quality and duration dimensions because the perception of speech involves the perception of spectral (e.g., pitch) and temporal information. For instance, the perception of English tense-lax vowel contrasts involves listeners' perception of spectral and temporal differences between the vowels.

The stimuli in this task were the same as the stimuli used by Safronova and Mora (2013), which were seven Catalan vowels produced in isolation by two phonetically trained male and female Catalan native speakers whose voices (male vs. female) represented the quality dimension (i.e. the first task set). The vowel stimuli were digitally recorded (Marantz PMD660) in a soundproof booth at a 44.1 kHz sampling rate frequency with a 16-bit resolution and were normalized for amplitude (70 dB). To create the second task set (i.e. duration dimension), vowel stimuli were lengthened to 500 ms and shortened to 200 ms using the PSOLA algorithm in Praat. The vowel stimuli were then organized in a test block consisting of 224 randomly presented trials (7 vowels x 2 voices x 2 durations x 8 identical copies of each stimulus) with a break after first 112 trials. Eight warm-up trials were added at the beginning of the test block and another eight warm-up trials - after the break, which were excluded from the data analysis.

In each trial the presentation of the task sets was cued by two alternating pictures displayed in the center of the screen. The quality task (Q) was represented by a picture of male vs. female located on the left side of the screen and the duration task (D) was represented by a picture of long and short horizontal lines located on the right side of the screen (see Figure 13). The response labels were displayed below the pictures to remind participants about the key assignment. The picture corresponding to the upcoming task lit up blue and was followed by the presentation of a vowel stimulus. The pictures alternated predictably creating a sequence of repeat trials (i.e. QQ - DD - QQ ...), which required participants to stay on the same task, and shift trials (i.e. QD - DQ ...) requiring reallocation of their attention to the other task.

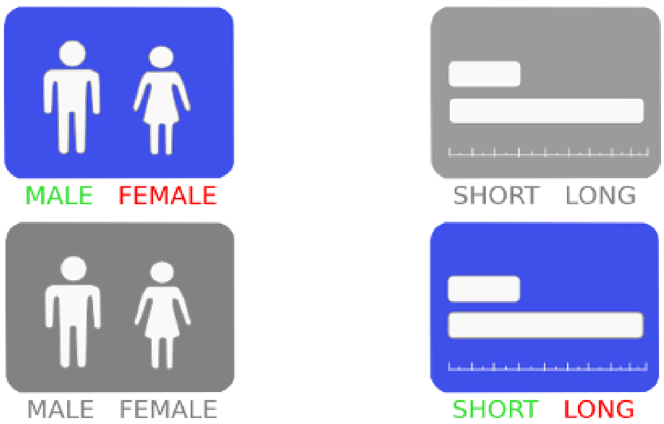


Figure 13. Visual cues to the attention-shifting task.

Participants were informed that the task consisted of two predictably alternating sub-tasks requiring identification of the voice gender or the duration of a vowel sound (see Appendix I, Figure I.1). They were instructed to follow the visual cues to the tasks and to decide as fast and accurately as possible whether the voice was male or female and whether the vowel was short or long by pressing the appropriate key. Prior to the test block participants completed a practice session consisting of tree practice blocks (48 trials in total). The first two blocks (12 trials each) provided separate training on the quality and duration sub-tasks. In the third block (24 trials) two task sets appeared in alternating runs, thus, simulating the test block procedure. In the practice blocks, if participants made a mistake or exceeded the 3000 ms response time limit the message "Oops!" and "Too slow!" were displayed on the screen, respectively. At the end of each practice and test block participants were presented with feedback showing their average error rate and reaction time. No feedback was provided during the test block in order to prevent participants from focusing on accuracy only.

The task was administered using OpenSesame software. The reaction times on trials where an incorrect response was given were discarded. For each participant the data were screened and the reaction times which were beyond 2.5 SD from the participant's mean reaction time for each sub-task and trial type (i.e. reaction time on quality/duration on repeat/shift trials) were excluded from the analysis (for a total of 3% of the data). The measures of participants' attention control were obtained by computing the response latency differences between shift and repeat trials, referred to as shift cost as well as mean error rate on repeat and shift trials (Prior & MacWhinney, 2010; Rogers & Monsell, 1995; Segalowitz & Frenkiel-Fishman, 2005). Regarding the accuracy measure, the total percentage of errors rather than the accuracy shift cost (Rogers & Monsell, 1995) was used due to non-significant differences between error rates in shift and repeat trials found in the present study (for details see Chapter 4). It was assumed that a more efficient phonological attention control, as indexed by lower shift cost and error rate scores, would be associated with learners' better ability to perceive cross-language phonetic distance and their higher level of L2 perceptual phonological competence.

4.2.6 Vocabulary size task

A vocabulary size estimate was used as a method to assess participants' proficiency level in English. Recent research on the role of vocabulary size in L2 speech perception has shown that a larger vocabulary size is associated with a more accurate L2 vowel perception (Bungaard-Nielsen et al., 2011). Therefore, in the present study participants' overall proficiency was measured through a receptive vocabulary score obtained from combining X_Lex (Meara, 2005) and Y_Lex test scores (Meara & Miralpeix, 2006) which estimate the vocabulary size in the 0-10000 word range. This scoring method has been found to reflect different L2 proficiency levels and has been widely used in previous L2 speech research as a reliable phonologically-related measure of L2 proficiency (Cerviño-Povedano & Mora, 2015; Darcy et al., 2014; Miralpeix, 2012).

In X_Lex and Y_Lex tests participants were presented with a series of English words appearing in the middle of the screen and were asked to decide whether or not they knew the meaning of each word by clicking the appropriate button (a *happy* or a *sad* face). Participants were asked to be honest and were informed that some of the words were English nonwords. The sum of X_Lex and Y_Lex corrected scores was used as a measure of the participants' proficiency level[7].

4.3 General procedure

The L2 learners were tested in small groups of two participants in a quiet room at the University of Barcelona (Spain). The experiment was conducted in a single session that lasted approximately one hour and a half. The tasks were administered using two PCs and two sets of soundproof headphones to ensure maximum sound quality and to avoid any external distractions. Taking into account possible effects of the task order on the re-

[7] There are two types of X_Lex and Y_Lex test scores: raw scores and corrected scores. The latter are the raw scores adjusted to the number of nonwords the participants claimed they knew.

sults of the experiment, the task presentation followed one of two orders counterbalanced across participants (see Table 6).

Table 6
General procedure: tasks order.

Order 1	**Order 2**
Attention shifting task[a]	Serial nonword recognition task[c]
Target sound/Serial sound recognition task[b]	Perceptual assimilation task[e]
Vowel discrimination task (ABX)[d]	Serial sound/Target sound recognition task
Vocabulary size task (X_Lex and Y_Lex)[f]	

This task distribution resulted in half of the participants beginning the experiment with the acoustic short-term memory task (either serial sound recognition task or the target sound recognition task) and the other half of the participants doing phonological short-term memory first. The acoustic short-term memory tasks were never presented sequentially or immediately before or after the phonological short-term memory task. Once a participant completed the first task set (Order 1) he/she was asked to switch to another PC in order to complete the second task set (Order 2) and vice versa.

The group of native speakers of American English was tested individually under sound-attenuated conditions in the L2 Psycholinguistics lab at Indiana University (Bloomington, USA). The order of the tasks was the same for all participants in the group. They were first asked to fill out the linguistic background questionnaire and then performed the ABX vowel discrimination task, which provided baseline data in the present study.

Chapter 5

Results

This chapter provides a detailed description of the data analyses and reports the results obtained regarding the relationship between the L2 learners' cognitive abilities, perception of cross-language phonetic distance and L2 perceptual phonological competence. The first part of the chapter is devoted to the results of preliminary analyses of the L2 learners' performance in L2 vowel perception and cognitive tasks. Specifically, the description of the participants' performance in the perceptual assimilation task and the L2 vowel discrimination task is provided. The L2 learners' L2 vowel discrimination scores are compared with the baseline data. The results of exploratory analyses aimed at validating the work of the cognitive tasks are also reported in the first part of this chapter.

The second part of the chapter focuses on the contribution of phonological short-term memory, acoustic short-term memory and attention control to L2 learners' acquisition of perceptual phonological competence. The results regarding the contribution of factors such as age of onset of L2 learning, L2 proficiency, length of stay abroad and amount of L2 use to learners' perception of cross-language phonetic distance and to their L2 perceptual phonological competence are also reported. The research questions are answered and the predicted hypothesis is verified by reporting the results of a series of correlation analyses between the measures of L2 learners' cognitive skills and L2 vowel perception. Prior to performing correlation analyses exploratory data analyses were conducted for all the measures obtained to verify the normality of distribution and homogeneity of variance needed to perform parametric tests (Field, 2009). All statistical analyses were performed in SPSS 20. An alpha level of .05 was used as a significance criterion in the present study.

5.1 L2 vowel perception

The present section reports the results of the cross-language perceptual assimilation task and the categorical L2 vowel discrimination task which were used to assess the participants' L2 vowel perception. The assimilation scores of Catalan and English vowels are presented. The focus is made on the perceived cross-language phonetic similarity and the cross-language assimilation overlap scores measuring the participants' degree of perceived phonetic distance between L2 and L1 vowels, and the degree of perceptual overlap between contrasting L2 vowels, respectively. The section further reports the results of the categorical L2 vowel discrimination task. The learners' L2 vowel discrimination scores are compared with those of the native speakers.

5.1.1 Perception of cross-language phonetic distance

This section reports the results of the perceptual assimilation task. The identification scores of the control Catalan (/i/, /e/, /ɛ/ and /a/) and the test English (/i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/) vowels are presented. The section also describes the perceived cross-language phonetic similarity and the cross-language assimilation overlap scores which measured the participants' perception of a cross-language phonetic distance between L2 and L1 vowels.

5.1.1.1 Identification of Catalan vowels

The average percentages of identification of the control Catalan vowel stimuli in terms of Catalan vowel categories are presented in Table 7. The average goodness-of-fit ratings (in parentheses) indicate the average ratings of the perceived degree of similarity between the Catalan vowels and the selected vowel response categories. The average fit indexes represent a combination of identification and goodness-of-fit scores (Guion et al., 2000). For example, to calculate a fit index for Catalan vowel /a/, the proportion of assimilation of this vowel to Catalan /a/ was multiplied by the goodness-of-fit rating to obtain a score out of seven (e.g., 0.97 x 5.0 = 4.8), where the maximum score indicates the perfect fit between the vowel stimulus and the selected L1 vowel category.

Contrary to our expectations the control Catalan vowel stimuli were not identified at perfect 100% accuracy rate. As the data in Table 7 demonstrate, the correct identification rates were the highest for Catalan /a/ and /i/. The former was correctly identified by the participants 97% of the time and received an average goodness-of-fit rating of 5. Participants identified Catalan /i/ at 82% accuracy rate with an average goodness-of-fit rating of 4.5. Catalan /a/ and /i/ obtained the highest mean fit indexes of 4.8 and 3.7, respectively, which however did not indicate a perfect match to the native vowel categories. The results for Catalan /ɛ/, which was correctly identified in 59% of instances with 4.6 goodness-of-fit rating and in 31% of cases heard as /e/, coincide with the previous research findings demonstrating that Catalan listeners may to some extent mis-identify and have difficulties discriminating /e/ and /ɛ/ (Cebrian, 2006; Pallier et al., 1997). Catalan /e/ was correctly identified least often (50%) and received an average goodness-of-fit rating of 4.9. Surprisingly, in 44% of instances Catalan /e/ was identified as Catalan /i/, which as shown in Figure 6, is acoustically close to Catalan /i/ and English /ɪ/ and /ɛ/ (see Section 4.2.2.1). As suggested by the mean fit indexes of 2.5 and 2.8, Catalan /e/ and /ɛ/ were perceived as relatively distinct from the Catalan /e/ and /ɛ/ vowel categories.

The unexpectedly low correct identification scores for Catalan /e/ and /ɛ/ obtained in the present study might be explained by several factors. First, the participants' vowel identification scores might have been affected by the task design in which the presentation of the control Catalan stimuli were mixed with the test English stimuli. Second, participants could have had difficulties in distinguishing Catalan /e/ and /ɛ/ due to the fact that these two Catalan mid vowels might be undergoing a process of merging, espe-

Table 7

Mean percentages identification and goodness-of-fit ratings (in parentheses) of Catalan vowel stimuli in terms of Catalan vowel categories.

Catalan vowel stimuli	Catalan response category							Fit index[a]
	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/	
/i/	**82** (**4.5**)	17 (4.5)	1 (4.2)					3.7 (1.52)
/e/	44 (3.6)	**50** (**4.9**)	6 (3.6)					2.5 (2.10)
/ɛ/	5 (2.4)	31 (4.1)	**59** (**4.6**)	5 (2.4)				2.8 (1.65)
/a/	1 (1.8)	2 (3.5)	**97** (**5.0**)					4.8 (1.19)

Note. The goodness-of-fit ratings are based on a scale from 1= *totally different* to 7= *identical*. Boldfaced values indicate the modal identification response. Mean fit indexes (*SD* in parentheses) are provided for the modal identification responses.
[a] Fit indexes were derived from the proportion of identifications and goodness ratings (see the text).

cially in Barcelona where both Catalan and Spanish are widely spoken. In addition, the participants' use of Spanish might have had an effect on the participants' perception of Catalan mid vowels (for the participants' amount of daily use of Catalan, Spanish and English see Chapter 4)[1]. Finally, the differences in the participants' experience with English might have also affected the identification of L1 vowels (Cebrian, 2006). Nonetheless, the fact that the percentages of identification and goodness-of-fit ratings of the Catalan vowels were relatively high and no instances of the misuse of the response categories were observed may suggest that the participants understood the task and did it properly.

5.1.1.2 Identification of English vowels

The participants' mean percentages of assimilation of the test English vowels (/i/, /ɪ/, /ɛ/, /æ/, /ʌ/ and /ɑ/) to the Catalan vowel categories and the mean goodness-of-fit ratings indicating the perceived degree of similarity between the L2 and L1 vowels are presented in Table 8. The highest assimilation rates were obtained by English /i/, /æ/, /ʌ/ and /ɑ/. English /i/ was identified as Catalan /i/ 95% of the time with the goodness-

[1]There is no mid /e/ and /ɛ/ vowel contrast in Spanish which includes only /e/ vowel sound.

of-fit rating= 4.7. English /æ/, /ʌ/ and /ɑ/ were assimilated to Catalan /a/ vowel 83%, 95% and 97%, respectively, and received relatively high goodness ratings (i.e. 4.1, 4.3 and 4.7, respectively). These data suggest that English /i/ and /æ/, /ʌ/ and /ɑ/ were perceived as similar to their Catalan counterparts by the participants in the present study. Overall the mean assimilation scores obtained for English /i/, /æ/, /ʌ/ and /ɑ/ in the present study coincide with the perceptual mapping data for the Catalan and American English vowels reported by previous research (Rallo Fabra, 2005).

Table 8
Mean percent identification and goodness-of-fit rating (in parentheses) of English vowel stimuli in terms of Catalan vowel categories.

English vowel stimuli	**Catalan response category**						
	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
/i/	**95** **(4.7)**	5 (3.8)					
/ɪ/	**42** **(3.1)**	37 (3.7)	21 (4.0)				
/ɛ/		14 (3.8)	**51** **(4.3)**	35 (3.3)			
/æ/		2 (4.1)	15 (3.9)	**83** **(4.1)**			
/ʌ/		1 (3.3)	1 (3.7)	**95** **(4.3)**	3 (2.7)		
/ɑ/				**97** **(4.7)**	3 (2.9)		

Note. The goodness-of-fit ratings are based on a scale from (1= *totally different* to 7= *identical*. Boldfaced values indicate the modal identification response.

Similar to the perceptual assimilation results reported by previous studies (Cebrian, 2009; Rallo Fabra, 2005), Table 8 shows that neither English /ɪ/ nor /ɛ/ were clearly mapped onto a single Catalan vowel category. Most frequently participants assimilated English /ɪ/ to Catalan /i/ in 42% of instances with an average goodness-of-fit rating of 3.1. However, 37% and 21% of the time the vowel was identified as Catalan /e/ and /ɛ/ and received mean goodness ratings of 3.7 and 4.0, respectively. English /ɛ/ was most frequently identified as Catalan /ɛ/ in 51% of instances with an average goodness-of-fit rating of 4.3. The second perceptually closest L1 vowel category was /a/ to which

English /ɛ/ was assimilated 35% of the time with an average goodness-of-fit rating of 3.3.

The assimilation scores of the English vowel stimuli were further analyzed by calculating the overall fit to the Catalan vowel categories (i.e. fit index) individually for each participant (Guion et al., 2000). The fit indexes for the English vowel stimuli as well as mean proportions of identification and goodness-of-fit ratings are shown in Table 9. The fit indexes ranged from 1.2 (English /ɛ/ identified as Catalan /a/) to 4.6 (English /ɑ/ identified as Catalan /a/). The highest mean fit indexes were obtained by English /i/ ($M = 4.5, SD = 1.4$) identified as Catalan /i/ and /ɑ/, /ʌ/ and /æ/ ($M = 4.6$, $SD = 1.3$; $M = 4.0$, $SD = 1.1$; $M = 3.4$, $SD = 1.2$) identified as Catalan /a/. These results suggest that English /i/, /ɑ/, /ʌ/ and /æ/ were perceived as best matching the selected L1 vowel categories. The pairwise comparisons (Bonferroni corrected $p < .02$) of the fit indexes revealed statistically significant differences in fit indexes between /ɑ/ and /ʌ/ ($z = -3.07$, $p < .02$, $r = -.46$), /ɑ/ and /æ/ ($z = -4.54$, $p < .001$, $r = -.68$) and /æ/ and /ʌ/ ($z = -2.69$, $p < .02$, $r = -.40$). These results suggest that despite being assimilated to the same Catalan /a/ vowel, English /ɑ/, /ʌ/ and /æ/ differed in the degree of fit to that L1 vowel category, thus indicating *Category-Goodness* assimilation (Best, 1995; Best & Tyler, 2007).

Table 9

Mean proportion of identifications, goodness-of-fit ratings and fit indexes for English vowels (SD in parentheses).

English vowels	**Catalan vowel category**	**Proportion of identification**	**Goodness rating**	**Fit index**
/i/	/i/	0.95 (0.1)	4.7 (1.2)	4.5 (1.4)
/ɪ/	/i/	0.42 (0.3)	3.1 (1.1)	1.3 (1.2)
	/e/	0.37 (0.3)	3.7 (1.4)	1.4 (1.3)
/ɛ/	/ɛ/	0.51 (0.3)	4.3 (1.2)	2.2 (1.5)
	/a/	0.35 (0.3)	3.3 (1.2)	1.2 (1.0)
/æ/	/a/	0.83 (0.2)	4.1 (1.1)	3.4 (1.2)
/ʌ/	/a/	0.95 (0.1)	4.3 (1.7)	4.0 (1.1)
/ɑ/	/a/	0.97 (0.1)	4.7 (1.2)	4.6 (1.3)

Note. Identifications below 30% are not included.

The mean fit indexes obtained for English /ɪ/ and /ɛ/, most frequently identified as Catalan /i/ (M = 1.3, SD = 1.2) and /ɛ/ (M = 2.2, SD = 1.5), respectively, suggest that these two English vowels were perceived as distinct from the respective Catalan categories. Also these results show that English /ɪ/ and /ɛ/ were assimilated to two distinct L1 categories, predicting good discrimination of this L2 vowel contrast (Best, 1995; Best & Tyler, 2007). As Table 9 shows, both English /i/ and /ɪ/ were assimilated to the same Catalan /i/ vowel category. However, the difference in the fit indexes ($z = -5.83$, $p < .001$, $r = -.87$) may indicate *Category-Goodness* assimilation (Best, 1995; Best & Tyler, 2007).

The mean fit index of the English vowels to the Catalan vowel categories, calculated on the basis of the modal identification responses (see Table 8) was 3.3 ($SD = 0.7$). Similar mean fit index of 3.4 was obtained for the correctly identified control Catalan vowels ($SD = 0.8$). The results of the t-test showed that the difference in the overall fit indexes between Catalan and English vowels was not significant ($t(44) = .33$, $p > .05$), suggesting that both Catalan and English vowels were perceived as equally good instances of L1 Catalan vowel categories.

5.1.1.3 Perceived phonetic distance between L2 and L1 vowels

The participants' perceived phonetic similarity scores are presented in Table 10. As explained in Chapter 4, the perceived cross-language phonetic similarity scores were predicted to indicate the participants' degree of perceived phonetic distance between L2 and L1 vowels.

As Figure 14 shows, English /i/ and /ɑ/ were perceived the perceptually closest to the selected L1 categories as indicated by the mean perceived phonetic similarity scores of 65.6% ($SD = 18.8$) and 66.3% ($SD = 18.5$), respectively. English /æ/ and /ʌ/ obtained perceived phonetic similarity scores of 57.9% ($SD = 16.5$) and 50.0% ($SD = 15.7$). Finally, perceived phonetic similarity scores of English /ɪ/ and /ɛ/ ($M = 34.1\%$, $SD = 14.5$

and $M = 38.5\%$, $SD = 16.6$, respectively) may suggest that these English vowels were perceived relatively different from the selected L1 categories.

Table 10

Descriptive statistics for participants' perceived cross-language phonetic similarity scores (CLPS) in percentage.

English vowels	**Mean**	**SD**	**Median**	**Range**
/i/	65.6	18.8	68.6	18.6 - 92.9
/ɪ/	34.1	14.5	34.3	8.6 - 61.4
/ɛ/	38.5	16.6	35.7	12.9 - 75.7
/æ/	50.0	15.7	51.4	11.4 - 90.0
/ʌ/	57.9	16.5	54.3	22.9 - 90.0
/ɑ/	66.3	18.5	70.0	21.4 - 94.3
CLPS	52.0	10.2	52.8	30.7 - 74.0

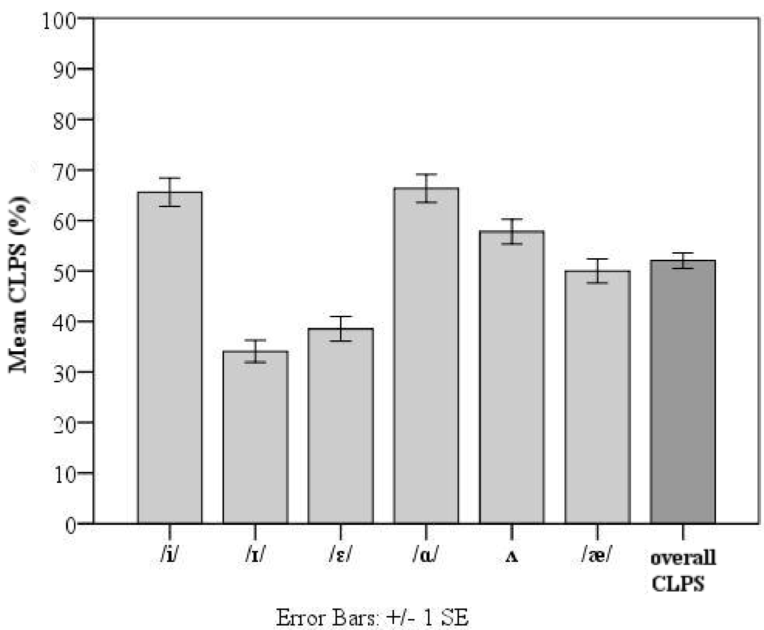


Figure 14. Participants' perceived cross-language phonetic similarity scores (CLPS).

The analyses conducted for English vowels showed that there was a significant effect of *Vowel* on the participants' perceived phonetic similarity scores ($\chi^2(5) = 93.11$, $p < .001$). As shown in Table 11, the pairwise comparisons (Bonferroni correction $p < .003$) revealed statistically significant differences between /i/ and /ɪ/, /i/ and /ɛ/, /i/ and /æ/, /ɛ/ and /ɑ/, /ɛ/ and /ʌ/, /ɛ/ and /æ/, /ɑ/ and /ʌ/, /ɑ/ and /æ/, /ɪ/ and /ɑ/, /ɪ/ and /ʌ/, /ɪ/ and /æ/. There were no statistically significant differences between /i/ and /ɑ/,

/i/ and /ʌ/, /ʌ/ and /æ/, and /ɪ/ and /ɛ/. These results suggest that vowels in L2 vowel contrasts differ in the degree of perceived similarity to the perceptually closest L1 vowel category. In other words some test L2 vowels were perceived as being perceptually closer to the L1 vowels than other L2 vowels.

Table 11
The results of the pairwise comparisons of cross-language phonetic similarity scores for each English vowel pair.

Vowels	*z*	*p*	*r*
/i/ vs. /ɪ/	-5.52	$< .001$	-.82
/i/ vs. /ɛ/	-5.13	$< .001$	-.76
/i/ vs. /æ/	-3.86	$< .001$	-.57
/ɛ/ vs. /ɑ/	-5.19	$< .001$	-.77
/ɛ/ vs. /ʌ/	-4.17	$< .001$	-.62
/ɛ/ vs. /æ/	-3.27	$< .003$	-.49
/ɑ/ vs. /ʌ/	-3.31	$< .003$	-.49
/ɑ/ vs. /æ/	-4.68	$< .001$	-.70
/ɪ/ vs. /ɑ/	-5.35	$< .001$	-.80
/ɪ/ vs. /ʌ/	-4.96	$< .001$	-.74
/ɪ/ vs. /æ/	-4.14	$< .001$	-.62
/i/ vs. /ɑ/	-.33	$> .05$	-.05
/i/ vs. /ʌ/	-2.77	$> .003$	-.41
/ʌ/ vs. /æ/	-2.38	$> .01$	-.35
/ɪ/ vs. /ɛ/	-1.52	$> .05$	-.23

5.1.1.4 Cross-language assimilation overlap

The mean assimilation overlap scores for English /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/- /æ/ and /ʌ/-/æ/ vowel contrasts are presented in Table 12 and Figure 15. As explained in Chapter 4, cross-language assimilation overlap score measured the degree of perceptual assimilation overlap between vowels in the above mentioned L2 vowel contrasts and served as a measure of distance between perceptually similar L2 vowels in the L2 learners' phonological system (Levy, 2009).

The data analyses revealed a statistically significant effect of *Vowel Contrast* on the participants' assimilation overlap scores ($\chi^2(4) = 119.12$, $p < .001$). As shown in Table 13, the pairwise comparisons (Bonferroni corrected $p < .005$) revealed statistically significant differences between English /i/-/ɪ/ and /ʌ/-/ɑ/, /i/-/ɪ/ and /ʌ/-/æ/, /i/-/ɪ/

Table 12

Descriptive statistics for participants' cross-language assimilation overlap scores (CLAO) in percentage.

English vowel contrast	Mean	SD	Median	Range
/i/ - /ɪ/	45.98	29.58	50.0	0.0 - 100
/ɪ/ - /ɛ/	29.57	19.01	25.8	0.0 - 75.0
/ɑ/ - /ʌ/	94.29	11.79	100.0	41.7 - 100
/ɑ/ - /æ/	81.63	20.69	90.9	25.0 - 100
/ʌ/ - /æ/	82.89	17.74	90.9	33.3 - 100
Total CLAO	66.87	10.34	70.0	45.3 - 81.7

and /ɑ/-/æ/, /ɪ/-/ɛ/ and /ʌ/-/ɑ/, /ɪ/-/ɛ/ and /ʌ/-/æ/, /ɪ/-/ɛ/ and /ɑ/-/æ/, /ʌ/-/æ/ and /ʌ/-/ɑ/, and /ɑ/-/æ/ and /ʌ/-/ɑ/. There were no statistically significant differences between English /i/-/ɪ/ and /ɪ/-/ɛ/, and between /ɑ/-/æ/ and /ʌ/-/æ/. The results suggest that L2 vowel contrasts differ in the degree to which the contrasting vowels are perceptually assimilated by the participants to the same L1 category. This means that some contrasting L2 vowels might be perceived as being phonetically closer to each other in comparison with other L2 vowel pairs.

Table 13

The results of the pairwise comparisons of cross-language assimilation overlap scores for English vowel pairs.

Vowel pairs	*z*	*p*	*r*
/i/-/ɪ/ vs. /ʌ/-/ɑ/	-5.48	$< .001$	-.82
/i/-/ɪ/ vs. /ʌ/-/æ/	-5.14	$< .001$	-.77
/i/-/ɪ/ vs. /ɑ/-/æ/	-4.91	$< .001$	-.73
/ɪ/-/ɛ/ vs. /ʌ/-/ɑ/	-5.84	$< .001$	-.87
/ɪ/-/ɛ/ vs. /ʌ/-/æ/	-5.76	$< .001$	-.86
/ɪ/-/ɛ/ vs. /ɑ/-/æ/	-5.73	$< .001$	-.85
/ʌ/-/æ/ vs. /ʌ/-/ɑ/	-3.80	$< .001$	-.57
/ɑ/-/æ/ vs. /ʌ/-/ɑ/	-3.84	$< .001$	-.57
/i/-/ɪ/ vs. /ɪ/-/ɛ/	-2.77	$> .005$	-.41
/ɑ/-/æ/ vs. /ʌ/-/æ/	-1.08	$> .05$	-.16

Taken together with the English vowel assimilation data presented in Table 9, the results reported above suggest that the target L2 vowel contrasts may differ in their degree of discriminability. For example, high discrimination accuracy may be predicted for English /ɪ/ and /ɛ/ which were most of the time mapped onto two distinct L1 cate-

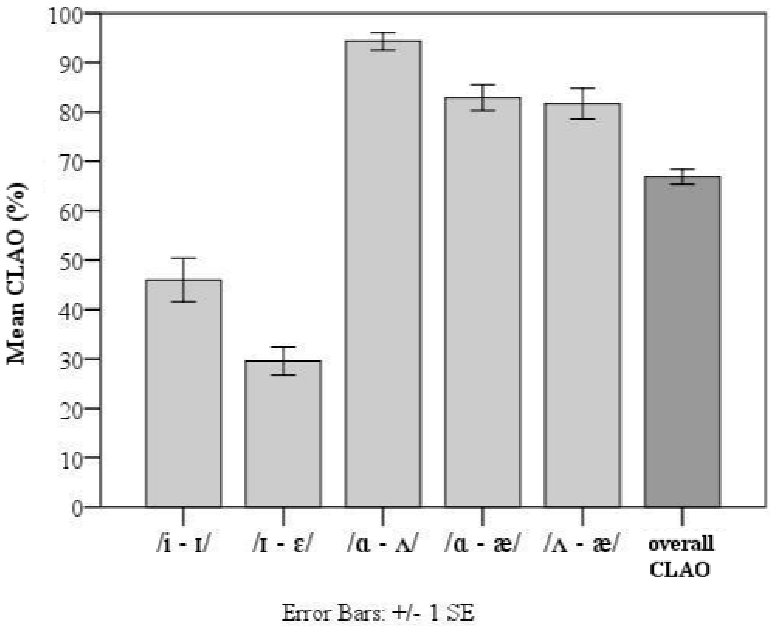


Figure 15. Participants' cross-language assimilation overlap scores (CLAO).

gories and showed a mean overlap of 29.57% (see Figure 16). Some lower discrimination accuracy rates may be predicted for English /i/-/ɪ/ which almost 50% of the time ($M = 45.98$) were assimilated to a single Catalan /i/ and differed in the goodness-of-fit ratings suggesting *Category-Goodness* assimilation (Best & Tyler, 2007). A rather low discriminability is expected for English /ɑ/-/æ/ and /ʌ/-/æ/ contrasts which on average perceptually overlapped in 81.63% and 82.89% of cases, respectively. The English vowel contrast /ɑ/-/ʌ/ is predicted to be difficult to discriminate as indicated by their fit indexes to a single Catalan /a/ vowel (see Table 9) as well as by the mean assimilation overlap of 94.29%.

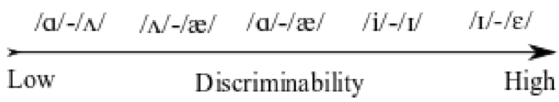


Figure 16. Prediction of discriminability of English vowel contrasts.

5.1.1.5 Inter-learner variation in perceived cross-language phonetic distance

Overall, the results obtained by the perceptual mapping task are in line with the data reported by previous cross-language studies (Cebrian, 2006; Rallo Fabra, 2005). The perceptual assimilation results of the present study show that the Catalan-Spanish bilingual EFL learners tended to assimilate English vowels to the perceptually closest Catalan counterparts (see Figure 6, Section 4.2.2.1). Moreover, participants were found to assimilate the contrasting L2 vowels to a single L1 category. However, as shown in Table 12 and Table 13, participants varied in the degree of perceived phonetic distance between L2 and L1 vowels ($M = 52.0$, $SD = 10.2$, $range = 30.7 - 74.0$) as well as between the contrasting L2 vowels ($M = 66.87$, $SD = 10.34$, $range = 45.3 - 81.7$). This inter-subject variation suggests that there were L2 learners who were likely to have established distinct phonetic categories for contrasting L2 vowels as reflected by the lower perceived phonetic similarity and assimilation overlap scores.

The Shapiro-Wilk's test revealed that the participants' overall perceived phonetic similarity scores were normally distributed ($p > .05$) whereas cross-language assimilation overlap scores violated the assumption of normality ($p < .05$). Therefore, both parametric and nonparametric tests were performed to examine the contribution of L2 learners' cognitive skills to the observed inter-subject variation in the degree of perceived phonetic distance between L2 and L1 vowels and the degree of perceptual overlap between L2 vowels. The results of the analyses are reported in the second part of this chapter.

5.1.2 Perception of L2 phonological contrasts

The present section reports the results regarding the participants' perception of English /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ vowel contrasts. In order to verify whether L2 learners' perceptual phonological competence reached native-like level, their L2 vowel discrimination scores were compared with those of the native American English speakers.

5.1.2.1 L2 vowel discrimination scores

Exploratory data analyses showed that the L2 learners' (L2Ls) and native speakers' (NSs) overall discrimination accuracy scores and discrimination reaction time scores were normally distributed as the results the Shapiro-Wilk's test were not significant ($p > .05$). The

Levene's test showed that the variances were equal for L2 learners and native speakers ($F(1,57) = 0.68, p > .05$) in the case of the overall discrimination reaction time scores, but for overall discrimination accuracy scores the variances were significantly different in the two groups ($F(1,57) = 18.65, p < .01$)[2].

The contrast-wise inspection of the data revealed that in general the discrimination accuracy scores were not normally distributed within both groups of participants ($p < .05$), except for L2 learners' and native speakers' correct discrimination accuracy scores of the /ɑ/-/ʌ/ contrast and the /i/-/ɪ/ contrast for L2 learners ($p > .05$). The Shapiro-Wilk's tests showed that the native speakers' reaction times did not violate the assumption of normality ($p > .05$). However, L2 learners' reaction times in all vowel contrasts, except for /i/-/ɪ/, were not normally distributed ($p < .05$). The Levene's test showed that the L2 learners' and native speakers' discrimination reaction time scores in each vowel contrast did not violate the assumption of homogeneity of variance ($p > .05$). However, in the case of discrimination accuracy scores, the variances were equal only for discrimination accuracy scores of the control /i/-/ɑ/ ($F(1,57) = 3.97, p > .05$) and the test /ɑ/-/ʌ/ ($F(1,57) = 3.45, p > .05$) vowel contrasts.

Taking these results into account, parametric tests were applied to assess L2 learners' overall performance on the L2 vowel discrimination task. The t-tests conducted for the inter-group comparisons in the case of overall discrimination accuracy scores, were adjusted for the violation of the assumption of homogeneity of variance. For the inter-group comparisons of discrimination accuracy scores for vowel pairs both nonparametric and parametric tests (in the case of identification scores for /ɑ/-/ʌ/ and reaction times in /i/-/i/) were applied.

[2]When comparing groups of native and non-native speakers data are likely to violate the assumption of homogeneity of variance where non-native speakers demonstrate a much larger variance than native speakers (Larson-Hall, 2009).

5.1.2.2 Comparing L2 learners' and native speakers' performance

Descriptive statistics for discrimination accuracy scores and discrimination reaction time scores obtained in the categorical ABX discrimination task are presented in Tables 14 and 15. There was a statistically significant effect of *Vowel Contrast* on the L2 learners' percentages of correct identification ($\chi^2(5) = 131.08$, $p < .001$) whereas for native speakers the effect was not significant ($\chi^2(5) = 10.13$, $p > .05$). Wilcoxon signed rank tests (Bonferroni corrected $p < .003$) conducted for the pairwise comparisons showed that for L2 learners there were statistically significant differences among all the L2 vowel contrasts ($p < .003$) except for /i/-/ɪ/ and /ɑ/-/æ/ ($p > .05$). In the case of native speakers there were no statistically significant differences in percentages of correct identification among the L2 vowel contrasts ($p > .003$).

There was also a significant effect of *Vowel Contrast* on L2 learners' reaction times ($\chi^2(5) = 98.55$, $p < .001$). The pairwise comparisons (Bonferroni corrected $p < .003$) showed that there were significant differences in the discrimination reaction time scores among all the L2 vowel contrasts except for /i/-/ɪ/ and /ɑ/-/ʌ/, /i/-/ɪ/ and /ʌ/-/æ/, /ʌ/-/æ/ and /ɑ/-/æ/, and /ʌ/-/æ/ and /ɑ/-/æ/ ($p > .05$). There was a statistically non-significant effect of *Vowel Contrasts* on the native speakers' discrimination reaction time scores (Wilks' Lambda = 0.63, $F(5,9) = 0.79$, $p > .05$, $\eta_p^2 = .31$). The pairwise comparisons showed that the difference among the native speakers' reaction times on L2 vowel contrasts were not significant ($p > .05$). These results suggest that native speakers were able to accurately and rapidly discriminate all test L2 vowels. In contrast, L2 learners' discrimination accuracy and speed seemed to vary as a function of the vowel contrast, suggesting that some L2 phonological contrasts were more difficult for them to acquire than others.

Table 14 and Figure 17 show that native speakers were overall more accurate in the L2 vowel discrimination task ($M = 96.3$, $SD = 2.2$, $range = 91.7 - 99.2$) than L2 learners ($M = 81.2$, $SD = 7.0$, $range = 65.8 - 93.3$). The independent-samples t-tests showed

Table 14
Average % DIS (SD in parentheses) for the L2 learners (L2Ls) and native speakers (NSs) groups.

Contrast	% DIS									
	L2Ls					NSs				
	M	*SD*	*SE*	*Mdn*	*Lowest–Highest*	*M*	*SD*	*SE*	*Mdn*	*Lowest–Highest*
/i/ - /a/	95.9	4.7	0.69	95.8	83.3-100	97.1	2.4	0.6	96.00	92.0-100
/i/ - /ɪ/	79.0	9.3	1.39	79.2	62.5-100	97.3	3.9	1.0	100.00	87.5-100
/ɪ/ - /ɛ/	91.4	6.7	1.00	91.7	79.2-100	96.7	3.7	1.0	95.83	87.5-100
/ɑ/ - /ʌ/	70.2	9.4	1.41	70.8	50.0-87.50	92.0	6.6	1.8	91.67	79.2-100
/ɑ/ - /æ/	79.4	14.1	2.10	83.3	41.7-100	97.0	3.8	1.0	97.92	87.5-100
/ʌ/ - /æ/	85.8	10.9	1.62	87.5	58.3-100	98.5	2.1	0.5	100.00	95.8-100
% DIS	81.2	7.0	1.0	82.5	65.8-93.3	96.3	2.2	0.6	96.7	91.7-99.2

that the difference between the groups was significant ($t(56.99) = -12.65$, $p < .001$) and presented a large effect size of $r = .86$. A series of Mann-Whitney U tests and an independent-samples t-test (in the case of English /ɑ/-/ʌ/) were conducted to compare L2 learners' and native speakers' accuracy scores for each vowel contrast. As expected, the analyses revealed statistically non-significant differences in discrimination of the control /i/-/ɑ/ contrast ($U = 242$, $z = -1.36$, $r = .18$). However, the tests showed that there were statistically significant inter-group differences in discrimination of each one of the test contrasts: /i/-/ɪ/ ($U = 29.50, z = -5.12, p < .001, r = .67$), /ɪ/-/ɛ/ ($U = 166.50$, $z = -2.71$, $p < .01$, $r = .35$), /ɑ/-/æ/ ($U = 57$, $z = -4.63$, $p < .001$, $r = .60$), /ʌ/-/æ/ ($U = 41.50$, $z = -4.95$, $p < .001$, $r = .64$) and /ɑ/-/ʌ/ ($t(57) = -8.01$, $p < .001$, $r = .73$).

As demonstrated by Table 15 and Figure 18, native speakers were also faster ($M = 396.9$, $SD = 133.5$) in the vowel discrimination task than L2 learners ($M = 540.3$, $SD = 172.9$). The independent-samples t-tests showed that the differences in reaction times between native speakers and L2 learners were significant ($t(57) = 2.84$, $p < .01$, $r = .35$). A series of Mann-Whitney U tests and an independent-samples t-test revealed statistically significant inter-group differences in the reaction times on test /i/-/ɪ/ ($t(52.09) = -10.59$, $p < .001$, $r = .83$), /ɑ/-/ʌ/ ($U = 182$, $z = -2.37$, $p < .05$, $r = .31$), /ɑ/-/æ/ ($U = 195$, $z = -2.14$, $p < .05$, $r = .28$) and /ʌ/-/æ/ ($U = 190$, $z = -2.23$, $p < .05$,

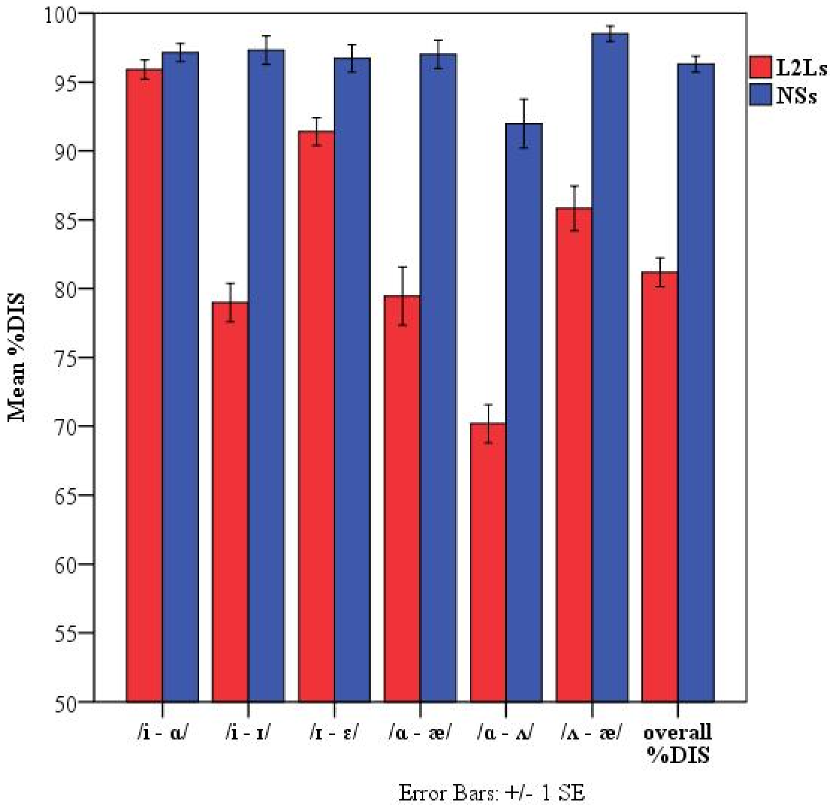


Figure 17. Native speakers' and L2 learners' percentage of correct discrimination (% DIS) in the L2 vowel discrimination task.

$r = .29$) contrasts. There were no statistically significant differences between native speakers' and L2 learners' reaction times on the control /i/-/ɑ/ and test /i/-/ɛ/ vowel contrasts ($U = 304$, $z = -.20$, $r = .03$ and $U = 239$, $z = -1.35$, $r = .18$, respectively).

Pearson product-moment correlation coefficients were computed to assess the relationship between the discrimination accuracy scores and discrimination reaction time scores. The correlation analysis of native speakers' and L2 learners' data revealed non-significant correlation between the two measures (NSs: $r = .485$, $N = 14$, $p(two-tailed) > .05$; L2Ls: $r = -.25$, $N = 45$, $p(two-tailed) > .05$). Therefore, faster reaction times on correctly identified trials were not associated with higher accuracy rate in the L2 vowel discrimination task. In other words, the participants who discriminated L2 vowels more accurately did not respond faster on the correct trials than the participants who performed less accurately on the vowel discrimination task. This suggests that accuracy and speed in discrimination measured two unrelated constructs and that the reaction times might have failed to measure individual differences in the efficiency of L2 vowel discrimination in the present study.

Table 15

Average reaction times (SD in parentheses) for the L2 learners (L2Ls) and native speakers (NSs) groups.

Contrast	Reaction time									
	L2Ls					NSs				
	M	*SD*	*SE*	*Mdn*	*Range*	*M*	*SD*	*SE*	*Mdn*	*Range*
/i/ - /a/	411.5	140.4	20.9	391.0	663	396.4	125.1	33.4	404.5	414
/i/ - /ɪ/	570.4	193.3	28.8	530.8	818	393.0	139.6	37.3	390.9	435
/ɪ/ - /ɛ/	446.1	151.9	22.6	409.5	680	383.0	127.0	33.9	369.1	389
/ɑ/ - /ʌ/	565.7	199.7	29.8	536.8	849	416.6	157.7	42.1	360.9	481
/ɑ/ - /æ/	503.5	176.0	26.2	496.4	903	394.8	119.5	31.9	431.2	395
/ʌ/ - /æ/	542.9	207.6	30.9	500.3	916	406.4	151.9	40.6	386.7	500
DIS RT	540.3	172.9	25.8	518.0	734	396.9	133.5	35.7	411.5	431

Overall, the analyses of baseline data suggest that the task worked as expected. Thus, no modifications to the measures of the L2 learners' L2 vowel discrimination ability were made. In sum, the analysis of the L2 vowel discrimination scores showed that native speakers significantly outperformed L2 learners in the ability to perceive vowel contrasts, which are phonological in the native speakers' but not in L2 learners' L1, demonstrating above 96% accuracy rate. Thus, a range of 90-100% total correct discrimination accuracy was set as a native-like level of vowel discrimination ability. The L2 learners performed the L2 vowel discrimination task at an average accuracy rate of 80%, which was below the native-like level. Overall, L2 learners were not only significantly less accurate but also much slower in the discrimination of L2 vowel contrasts than the group of native speakers. Moreover, the analysis of discrimination accuracy scores indicated a much bigger intra-group variation in accurate L2 vowel perception among L2 learners (SD = 7.0, $range = 65.8 - 93.3$) than among native speakers ($SD = 2.2$, $range = 91.7 - 99.2$). These results may suggest that there were L2 learners who were likely to establish L2 phonetic categories for the target L2 vowels and, thus, attain a more native-like level of L2 perceptual phonological competence. The following sections are devoted to the results regarding the contribution of the cognitive skills to the observed individual differences in L2 vowel perception.

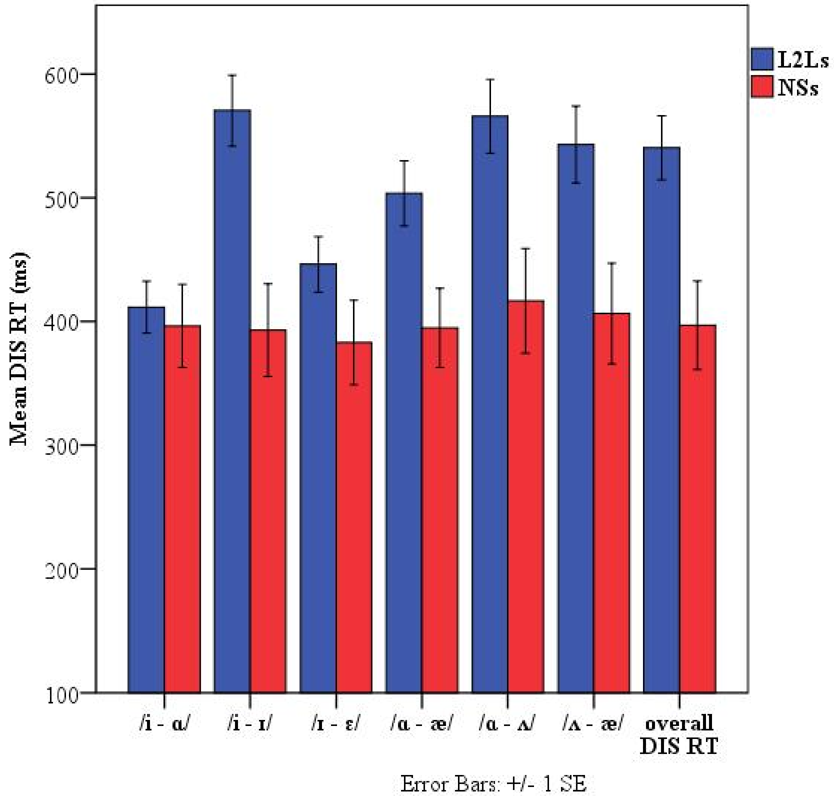


Figure 18. Native speakers' and L2 learners' average reaction times (RTs) in the vowel discrimination task.

5.1.3 Relationship between L2 speech perception measures

In order to examine the relationship between the measures used to assess the participants' L2 vowel perception, a series of *Pearson r* and *Spearman* r_s correlation analyses were conducted. As shown in Table 17, there was no significant correlation between the participants' perceived phonetic similarity scores, assimilation overlap scores, discrimination accuracy and discrimination reaction time scores ($p(two-tailed) > .05$). The results suggest that the degree of perceived phonetic distance between L2 and L1 vowels, perceptual overlap of L2 vowels and L2 vowel discrimination ability were unrelated. However, the correlation analysis between assimilation overlap scores and discrimination accuracy scores based on one-tailed distribution revealed a weak but significant correlation between the two measures of L2 speech perception ($r_s = -.256$, $p(one-tailed) < .05$). The negative correlation suggests that the lower overlap scores contributed to the higher accuracy in L2 vowel discrimination. Thus, the lower degree of perceptual overlap between contrasting L2 vowels may to some extent contribute to accurate perception of L2 vowel contrasts.

The results obtained may suggest that the L2 vowel perception tasks used in this study might have measured different abilities involved in L2 speech perception. The ability to perceive overall phonetic similarity between L2 and L1 sounds, as measured by the perceptual assimilation task, may not be directly employed in perception of L2 speech contrasts which might to a larger extent draw on the perceived distance between contrasting L2 sounds. In contrast to the cross-language similarity score, the perceptual assimilation overlap score might be better able to capture this ability and as result was found to contribute to the accuracy in discrimination of L2 sounds. This idea may be supported by the SLM and the PAM-L2 prediction that the learning of L2 phonological contrasts might require L2 learners' ability to discern phonetic differences between contrasting L2 sounds (Best & Tyler, 2007; Flege, 1995). This finding also fits well Levy's (2009) results showing the relationship between the degree of perceptual assimilation overlap and discriminability of contrasting L2 sounds.

5.2 L2 learning-related factors and L2 speech perception

In order to control for the potential contribution of factors such as age of onset of L2 learning, L2 proficiency, L2 use and length of stay abroad to the perception of cross-language phonetic distance and the discrimination of contrasting L2 sounds, we carried out a series of *Pearson r* and *Spearman* r_s correlation analyses. As we assumed L2 learning-related factors might be predictors of L2 phonological competence, one-tail correlations were performed[3].

The analyses yielded a non-significant relationship between L2 proficiency, age of onset of L2 learning, amount of L2 use, exposure to American English and the participants' perceived phonetic similarity and assimilation overlap scores (see Table 17). The absence

[3] A one-tailed test sets the level of significance ($p < .05$) in one tail of the distribution.

of the relationship between the amount of exposure to American English, the perceived phonetic similarity and the assimilation overlap scores suggests that the differences in the amount of exposure to American English did not affect the vowel judgments in the perceptual assimilation task. The participants' length of stay abroad was found to significantly correlate (negatively) with the assimilation overlap scores ($r = -.350$, $p < .01$), suggesting that a longer stay abroad was associated with a lower degree of perceptual assimilation overlap between L2 vowels. However, the length of stay abroad and the perceived phonetic similarity scores were found to be unrelated. Therefore, a longer stay in an L2-speaking country may facilitate the L2 learners' ability to avoid assimilating a pair of perceptually similar L2 vowels to the same L1 category, but may not significantly contribute to the perceived degree of phonetic distance between L2 and L1 vowels.

The L2 proficiency, age of onset of L2 learning, amount of L2 use, length of stay abroad and exposure to American English were found to be unrelated to the discrimination accuracy scores and the discrimination reaction time scores (see Table 17), suggesting that the participants' L2 vowel discrimination ability as assessed in the present study was independent of these L2 speech learning-related factors. The results regarding the relationship between the participants' exposure to American English and L2 vowel discrimination ability may also indicate no effect of the American English vowel stimuli used in the ABX discrimination task on the participants' discrimination scores. Overall, the results obtained indicate that learners' ability to discriminate contrasting L2 sounds was not associated with L2 proficiency and other L2 speech learning-related factors.

5.3 Cognitive abilities

5.3.1 Performance on the cognitive tasks

Prior to examining the relationship between cognitive abilities and L2 speech perception exploratory data analyses were conducted. The results of the Shapiro-Wilk's test indicated that the data obtained by the cognitive tasks were normally distributed ($p > .05$). The analysis of the standardized scores revealed no univariate outliers for any of the cognitive variables examined in the present study[4]. Hence, parametric tests were applied in order to explore the relationship among the participants' phonological short-term memory, acoustic short-term memory and attention control scores.

Another goal of the exploratory data analyses was to confirm that the cognitive tasks assessing the participants' phonological short-term memory, acoustic short-term memory and attention control worked properly. Figure 19 shows the participants' performance on the cognitive tasks. A Friedman test was carried out to see if there was an effect of increasing sequence length on the participants' performance on the phonological and the acoustic memory tasks. The results revealed a significant effect of increasing sequence length on the participants' percentage of correct identification in the phonological short-term memory task ($\chi^2(2) = 47.94$, $p < .001$) and acoustic short-term memory serial sound recognition task ($\chi^2(2) = 45.50$, $p < .001$). However, the effect was non-significant in the case of the acoustic short-term memory target sound recognition task ($\chi^2(2) = 18.50$, $p > .05$). The results suggest that increasing sequence length placed demands on the participants' phonological short-term memory and acoustic short-term memory capacity measured using a serial nonword/sound recognition procedure.

Post hoc analysis with Wilcoxon signed-rank tests was conducted for the pairwise comparisons. A Bonferroni correction was applied, resulting in a significance level set at $p < .017$. For the phonological short-term memory task median percentages of correct identification at five-, six-, and seven-item sequences were $Mdn = 87.5$, $Mdn = 62.5$ and $Mdn = 50.0$ (see Figure 19a). The analysis yielded statistically significant differences between the percentages of correct identification at five- and six- ($z = -5.05$, $p < .001$, $r = -.75$), five- and seven- ($z = -5.47$, $p < .001$, $r = -.81$) and six-and seven-item lengths ($z = -2.54$, $p < .017$, $r = -.36$).

[4]Cases with standardized scores in excess of ± 3.29 are potential outliers (Tabachnick & Fidell, 2013).

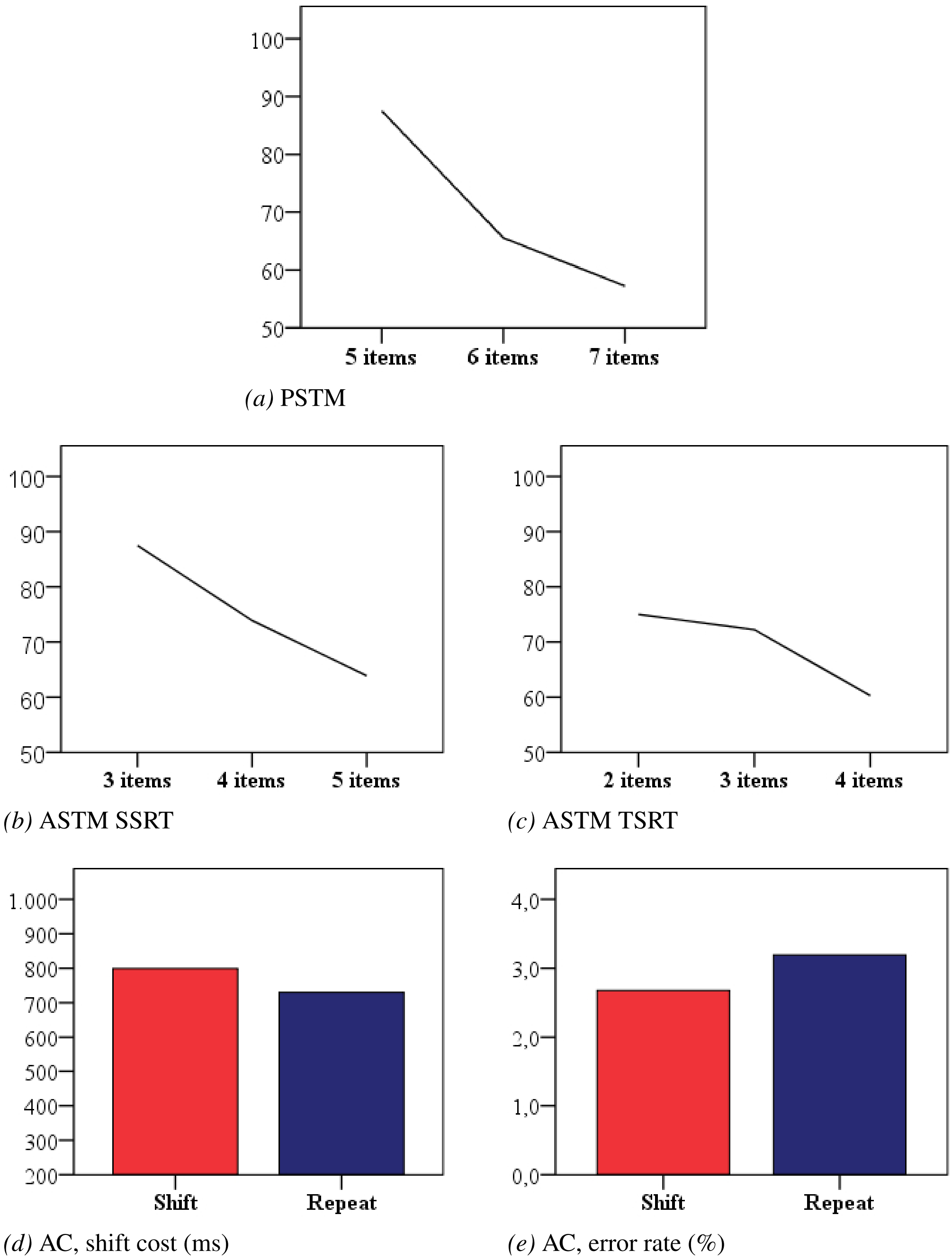


Figure 19. Median scores for the phonological short-term memory (PSTM), acoustic short-term memory serial sound recognition (ASTM SSRT) and target sound recognition (ASTM TSRT), and attention control (AC) tasks.

Regarding acoustic short-term memory as measured by the serial sound recognition task, median percentages of correct identification at three-, four-, and five-item sequences were $Mdn = 87.5$, $Mdn = 75.0$ and $Mdn = 62.5$ (see Figure 19b). Wilcoxon signed-rank tests also showed that the differences between the percentages of identification at three- and four- ($z = -4.62$, $p < .001$, $r = -.69$), three- and five- ($z = -5.36$, $p < .001$, $r = -.80$), and four- and five-item ($z = -3.04$, $p < .01$, $r = -.45$) were statistically significant. The analysis of the acoustic memory target sound recognition scores (see Figure 19c) revealed statistically significant differences between the percentages of correct identification at two- and four- ($Mdn = 75.0$, $Mdn = 62.5$, $z = -3.15$, $p < .01$, $r = -.47$), and three- and four-item lengths ($Mdn = 75.0$, $Mdn = 62.5$, $z = -3.78$, $p < .001$, $r = -.56$). However, there were no significant differences between the percentages of correct responses given on two- and three-item sequences ($Mdn = 75.0$, $Mdn = 75.0$, $z = -.94$, $p > .017$, $r = -.14$) indicating the same degree of difficulty in holding two- and three-item sequences in acoustic short-term memory.

Figures 19d and 19e illustrate the participants' average reaction time (RT) and error rate (ErR) in repeat and shift trials in the attention control task. The results of the Wilcoxon Signed Rank tests showed that as expected participants were significantly faster in repeat than in shift trials ($Mdn_{\text{repeat}} = 722.5$, $Mdn_{\text{shift}} = 798.8$, $z = -5.84$, $p < .001$, $r = -.87$). However, no statistically significant differences in error rates were found between shift and repeat trials ($Mdn_{\text{repeat}} = 3.6$, $Mdn_{\text{shift}} = 2.7$, $z = -1.04$, $p > .05$, $r = -.15$). These results may be explained by the participants' overall high accuracy rate in the attention control task ($M_{\text{ErR}} = 3.0$, $SD = 1.6$, $Mdn_{\text{ErR}} = 2.7$, $range = 0 - 7.1$). Taken together with the fact that there was no correlation between error rate and mean reaction time scores ($r = -.250$, $p > .05$) as well as between error rate and shift cost scores (SC, $r = .132$, $p > .05$) the results suggest that participants did not sacrifice accuracy for speed when staying on the same sub-task and shifting attention from one sub-task to the other.

A closer inspection of the attention control scores on quality and duration sub-tasks showed that participants were overall faster ($Mdn_{\text{RT}} = 892$, $Mdn_{\text{RT}} = 925$, $z = -4.95$, $p < .001$, $r = -.74$) and more accurate ($Mdn_{\text{ErR}} = 1.8$, $Mdn_{\text{ErR}} = 3.6$, $z = -3.89$, $p < .001$, $r = -.58$) when responding on quality than on duration trials. The participants were also found to respond much faster ($Mdn = 783$, $Mdn = 819$, $z = 4.05$, $p < .001$) when shifting attention to voice quality trials than to duration trials. These results suggest that the quality and duration dimensions in the attention control task might not be equal. However, no statistically significant differences were found between shift costs calculated separately for each task dimension ($Mdn_{\text{duration}} = 78.5$, $Mdn_{\text{quality}} = 94.2$, $z = -1.78$, $p > .05$, $r = -.27$), which indicates a similar degree of attentional cost for both task dimensions. Therefore, the overall attention control shift cost scores were used as a measure of the participants' attention control in further analyses. Overall, the results of the preliminary analyses suggest that, except for the acoustic memory target sound recognition task, the tasks assessing the participants' cognitive skills worked as expected.

5.3.2 Relationship between cognitive abilities

In order to explore the relationship between cognitive abilities, we conducted a series of *Pearson r* correlation analyses between the measures of phonological short-term memory, acoustic short-term memory and attention control. As Table 17 shows, there was no relationship between phonological memory scores, acoustic memory (either serial sound or target sound recognition scores) and attention control (either error rate or shift cost scores). The analyses revealed no relationship between the acoustic memory serial sound recognition task and the target sound recognition scores ($r = .243$, $p(two-tailed) > .05$). On the other hand, phonological short-term memory was found to correlate significantly with acoustic memory as assessed by the serial sound recognition task ($r = .438$, $p < .01$). However, a non-significant correlation was found between phonological short-term memory and acoustic short-term memory measured through the target sound recognition task ($r = .272$, $p(two-tailed) > .05$).

The results obtained show that there was some relationship between cognitive abilities. Specifically, the significant correlation between acoustic short-term memory (measured by the serial sound recognition task) and phonological short-term memory task indicate the existence of a relationship between phonological short-term memory and acoustic short-term memory which are thought to play in concert in speech processing (Strange, 2002; Wode, 1994). On the one hand, larger phonological short-term memory capacity may facilitate storage of larger amount of acoustic properties of speech sounds. On the other hand, in this study participants might have relied on both their phonological short-term memory and acoustic short-term memory in order to discriminate the sequences of CVC nonwords in the phonological short-term memory task. The observed correlations between the phonological memory scores and the acoustic memory serial sound recognition scores may also reflect the effects of the common task procedure. As discussed in previous research, serial order recognition and single item recognition may involve distinct processes (Henson, Hartley, Burgess, Hitch, & Flude, 2003). Thus, the similarities/differences in the task procedures may have caused a lack of relationship between the two acoustic short-term memory tasks and a moderate correlation between acoustic memory serial sound recognition scores and phonological memory scores.

The analyses of the relationship between the cognitive and L2 learning-related variables revealed that the participants' L2 proficiency level, as measured by the receptive vocabulary size, was negatively and weakly correlated with phonological short-term memory scores ($r = -.327$, $p(two-tailed) < .05$)[5]. This finding is consistent with previous research outcomes demonstrating that phonological short-term memory ability might be a more significant predictor of L2 proficiency and accurate pronunciation for lower proficiency L2 learners than for advanced L2 learners (Hummel, 2009). No relationship was

[5]L2 proficiency scores and percentage of exposure to American English were normally distributed whereas age of onset of L2 learning and L2 use did not meet the assumptions of normality of distribution. Nonparametric tests were applied to length of stay abroad scores since it could not be considered an interval measure.

found between acoustic short-term memory, attention control and any of the L2 learning-related factors (see Table 17).

Overall, the preliminary analyses suggest that the cognitive tasks used in the present study measured different constructs. The correlation found between phonological short-term memory and acoustic short-term memory scores indicates a relationship between the two types of short-term memory. On the other hand, these results may reflect some common cognitive mechanisms required in both tasks, such as an ability to hold sequences of speech-like stimuli in memory by means of subvocal rehearsal.

5.4 Cognitive abilities and L2 speech perception

The aim of this section is to address the central research question regarding the relationship between cognitive abilities and learners' acquisition of L2 perceptual phonological competence. The relationship was explored by means of a series of *Pearson r* and *Spearman* r_s zero-order correlation analyses. Since the present study aimed to examine the contribution of cognitive abilities to L2 phonological competence, the correlation analyses were based on one-tailed distributions. Table 16 presents the descriptive statistics for all the variables examined in the present study.

5.4.1 Cognitive abilities and perceived cross-language phonetic distance

In order to examine the relationship between cognitive abilities and participants' perception of cross-language phonetic distance, we correlated the measures of phonological memory, acoustic memory and attention control with the perceived phonetic similarity scores and the perceptual assimilation overlap scores. As Table 17 shows phonological short-term memory and acoustic short-term memory, as measured by the serial sound recognition task, significantly correlated with perceived phonetic similarity ($r = .312$, $p < .05$ and $r = .271$, $p < .05$, respectively). Greater phonological short-term memory and acoustic short-term memory capacity were associated with lower degree of perceived distance between L2 and L1 vowels. These results may suggest that both acoustic and phonological short-term memory may be involved in learners' categorization of L2 sounds as instances of perceptually close L1 sounds. Acoustic short-term memory operationalized as an ability to retrieve a single acoustic element from a large sequence of acoustic elements (i.e. as measured by the target sound recognition task) was found to be unrelated ($r = .175$, $p > .05$) to the degree of perceived cross-language phonetic distance. Therefore, short-term memory storages for the sequences of verbal and acoustic information rather than the ability to retrieve a particular acoustic element from speech input is associated with learners' ability to perceive phonetic distance between L2 and L1 sounds.

Table 16

Descriptive statistics for L2 perception, cognitive and demographic variables ($N = 45$).

Measure	**Mean**	**SD**	**Median**	**Lowest**	**Highest**
CLPS	52.0	10.2	52.8	30.7	74.0
CLAO*	66.9	10.3	70.0	45.3	81.7
% DIS	81.2	7.0	82.5	65.8	93.3
DIS RT	540.3	172.9	518	270	1004
ASTM SSRT	73.3	11.6	76.0	37.5	100
ASTM TSRT	66.4	8.9	66.7	51.4	87.5
PSTM	67.5	11.0	68.7	47.2	91.0
AC SC	96.3	59.3	93.8	3	222
AC ErR	3.0	1.6	2.7	0	7.1
L2 proficiency	6350	932.8	6350	4000	8500
AOL	6.2	2.2	6.0	2	15
L2 use	22.6	14.0	20.0	3.0	57.0
AmE exposure	47.8	26.2	50.0	0	100
LoS	3.3	7.6	0.0	0	40

Note. CLPS = perceived cross-language phonetic similarity; CLAO = cross-language assimilation overlap; % DIS = L2 vowel discrimination accuracy; DIS RT = L2 vowel discrimination reaction time; ASTM SSRT = acoustic short-term memory serial sound recognition task; ASTM TSRT = acoustic short-term memory target sound recognition task; PSTM = phonological short-term memory; AC SC= attention control shift cost; AC ErR = attention control error rate; L2 proficiency = combined X-Lex/Y-Lex vocabulary size score; AOL = age of onset of L2 learning; L2 use= hours/week of L2 use at home, work, with friends and reading, listening to music, etc.; AmE use = percentage of overall exposure to American English; LoS = length of stay abroad (weeks).

The attention control shift cost scores were also found to significantly correlate with the perceived phonetic similarity scores ($r = .510$, $p < .01$). In contrast to phonological short-term memory and acoustic short-term memory; higher attention control ability, as reflected by lower shift costs, was associated with higher degree of perceived phonetic distance between L2 and L1 vowels. Thus, a more efficient phonological attention control, which implies shifting attention between the acoustic cues, may promote a better ability to perceive phonetic distances between L2 and L1 sounds and facilitate the establishment of new phonetic categories for L2 sounds. The attention control error rate scores, which measured the participants' ability to accurately foreground the relevant acoustic dimension and inhibit the irrelevant one, were not found to be associated with the degree of perceived cross-language phonetic distance ($p > .05$).

The results of the correlation analyses between the cognitive abilities and the degree of L2 vowel perceptual assimilation overlap revealed a significant correlation between acoustic short-term memory serial sound recognition scores and assimilation overlap scores ($r_s = -.318$, $p < .05$). A larger acoustic short-term memory capacity was associated with a lower degree of the perceptual assimilation overlap between L2 vowels. Therefore, a larger acoustic short-term memory capacity may facilitate the establishment of distinct L2 vowel categories as reflected by a lower degree of single-category assimilation of contrasting L2 vowels. A statistically non-significant relationship (see Table 17) was found between phonological short-term memory, acoustic short-term memory (as measured by the target sound recognition task), attention control and assimilation overlap scores ($p > .05$), suggesting that these abilities might not be directly related to the degree of perceptual overlap between L2 vowels sounds assimilated to a single L1 category.

5.4.2 Cognitive abilities and L2 vowel discrimination

In order to explore the relationship between cognitive abilities and learners' perception of L2 vowel contrasts, we correlated the measures of phonological memory, acoustic memory and attention control with the L2 vowel discrimination accuracy and reaction time scores. As Table 17 shows, phonological short-term memory and acoustic short-term memory (as measured by the serial sound recognition task) correlated significantly with

Table 17

Simple correlations between perceptual, cognitive and demographic variables.

	Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	CLPS	1													
2	CLAO[a]	.215	1												
3	% DIS	.194	-.256	1											
4	DIS RT	-.152	-.190	-.249	1										
5	ASTM SSRT	**.271***[b]	**-.318***[b]	**.299***[b]	-.083	1									
6	ASTM TSRT	.175	-.168	.152	-.044	.250	1								
7	PSTM SNWR	**.312***[b]	-.182	**.291***[b]	-.226	**.438****	.271	1							
8	AC SC	**.510****[b]	-.026	.018	.170	.280	-.020	.096	1						
9	AC ErR	.245	.109	**.314***[b]	**-.317***[b]	.170	-.137	.237	.235	1					
10	L2 proficiency	.007	.019	-.016	-044	-.150	.048	**-.327***	-.213	-.008	1				
11	AOL[a]	-.099	-.052	-.131	.145	-.132	.020	-.035	-.138	**-.355***	-.100	1			
12	L2 use[a]	-.095	-.184	.037	.068	-.012	.129	-.013	-.086	.043	.079	-.181	1		
13	AmE exposure	.139	.146	.059	-.209	-.116	.165	-.070	-.194	-.025	**.313***	-.027	.182	1	
14	LoS[a]	.010	**-.350****[b]	-.082	.188	.143	.021	-.271	.100	-.101	.077	.042	.096	-.020	1

Note. $*p < .05$; $**p < .01$; $***p < .001$; CLPS = perceived cross-language phonetic similarity; CLAO = cross-language assimilation overlap; % DIS = L2 vowel discrimination accuracy; DIS RT = L2 vowel discrimination reaction time; ASTM SSRT = acoustic short-term memory serial sound recognition task; ASTM TSRT = acoustic short-term memory target sound recognition task; PSTM = phonological short-term memory; AC SC = attention control shift cost; L2 proficiency = vocabulary size estimate; AOL = age of onset of L2 learning; L2 use = hours/week of L2 use at home, work, with friends and reading, listening to music, etc.; AmE use = percentage of overall exposure to American English in comparrison with exposure to British English; LoS = length of stay abroad.

[a] *Spearman* correlation coefficients are presented.

[b] Based on one-tailed distributions.

discrimination accuracy scores ($r = .291, p < .05$ and $r = .299, p < .05$, respectively). These results suggest that larger phonological short-term memory and acoustic short-term memory capacities were associated with L2 learners' ability to accurately discriminate contrasting L2 vowels. Therefore, both types of short-term memory may facilitate the establishment of L2 phonetic categories.

As regards attention control, contrary to our expectations, attention control error rate but not attention control shift cost scores correlated significantly with the discrimination accuracy scores ($r = .314, p < .05$), that is, higher error rates in the attention control task were associated with a higher accuracy in the L2 vowel discrimination task. In other words, L2 learners with a lower ability to accurately foreground the relevant acoustic dimension were better able to accurately discriminate L2 vowels. These results may be explained by the fact that when discriminating L2 vowel contrasts participants might have over-relied on temporal differences between L2 vowels. Therefore, more accurate L2 vowel discrimination ability, as measured in the present study might be partly predicted by the participants' lower ability to inhibit irrelevant acoustic cues.

Attention control, as measured by error rate scores, was found to be the only variable which significantly correlated with discrimination reaction time scores ($r = -.317$, $p < .05$). This result suggests that less accurate attention-shifting was associated with faster reaction times on correctly identified trials in the L2 vowel discrimination task. In other words, accurately bringing relevant acoustic cues to the perceptual foreground may slow down accurate discrimination of contrasting L2 vowel sounds.

Moreover, having explored the relationship between phonological short-term memory, acoustic short-term memory, attention control and the participants' L2 vowel discrimination ability separately by contrast, we found that cognitive skills predicted the discrimination accuracy for some of the L2 contrasts and not for others. Specifically, phonological short-term memory ($r = .279$, $p < .05$), acoustic short-term memory ($r = .293$, $p < .05$) and attention control (ErR, $r = .298$, $p < .05$) contributed significantly to learners' accurate perception of English /ɑ/-/ʌ/ contrast. The acoustic short-term memory capacity was also found to be related to the discrimination of English /æ/ and /ʌ/ ($r_s = .326$, $p < .05$). These results indicate that the perception of English /ɑ/-/ʌ/ and /æ/-/ʌ/ contrasts was especially difficult and cognitively demanding for L2 learners as they perceptually assimilate these contrasting L2 vowels to the same native /a/ vowel (see Table 8).

5.5 Contribution of cognitive abilities to L2 speech perception

This section aims to answer two specific research questions asked in this dissertation:

1. To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' degree of perceived phonetic distance between L2 and L1 vowels?

2. To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' perception of L2 vowel contrasts?

A series of standard multiple regression analyses were conducted in order to estimate the unique contribution of phonological short-term memory, acoustic short-term memory and attention control to L2 learners' perception of cross-language phonetic distance and their perception of L2 vowel contrasts (see Table 18 and Table 19). This section reports the results obtained.

5.5.1 Contribution of cognitive abilities to cross-language vowel perception

This section answers the first research question regarding the extent to which phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' perception of phonetic distance between L2 and L1 vowels. Standard multiple regression analyses were conducted with the measures of cognitive abilities as predictor

Table 18

Significance of the regression models with the measures of cognitive abilities (Predictors) and L2 speech perception (Dependent variables).

Regression model	Dependent variables	Predictors	*F*	R^2	*p*	% Contribution
1	CLPS	PSTM ASTM SSRT AC SC	6.76	.331	$= .001$	**33.1%**
2	CLAO	ASTM SSRT LoS	5.72	.214	$< .01$	**21.4%**
3	% DIS	PSTM ASTM SSRT AC ErR	2.95	.178	$< .05$	**17.8%**

Note. CLPS = perceived cross-language phonetic similarity; CLAO = cross-language assimilation overlap; % DIS = L2 vowel discrimination accuracy; ASTM SSRT = acoustic short-term memory measured by the serial sound recognition task; PSTM = phonological short-term memory; AC SC = attention control shift cost; AC ErR = attention control error rate; LoS = length of stay abroad.

variables, and the perceived cross-language phonetic similarity scores and the perceptual assimilation overlap scores as dependent variables[6].

5.5.1.1 *Cognitive abilities and perceived cross-language phonetic distance*

In order to measure the extent to which cognitive abilities contribute to L2 learners' degree of perceived distance between L2 and L1 vowels, the regression model included the measures of phonological short-term memory, acoustic short-term memory (as measured by the serial sound recognition task) and attention control (as measured by shift cost) as predictor variables and the perceived phonetic similarity scores as a dependent variable (see Regression model 1 in Table 18). As Table 18 shows, the model was found to be statistically significant ($F(3,41) = 6.76$, $p = .001$) and accounted for 33.1% of variance in the perceived phonetic similarity scores ($R^2 = .331$, $AdjustedR^2 = .282$). As shown in Table 19, attention control was found to make the strongest unique 21.1% contribution ($beta = .479$, $Partr = .459$, $p = .001$), when the variance explained by phonological

[6]Prior to performing all standard multiple regression tests the assumptions of multicollinearity, normality, linearity and homoscedasticity had been tested (Tabachnick & Fidell, 2013). The results revealed no violations of the assumptions required for conducting linear regression.

Table 19

Estimated unique contribution of cognitive abilities (Predictors) to L2 speech perception (Dependent variables).

Regression model	**Dependent variables**	**Predictors**	***beta***	***Part r***	***p***	**% Unique contribution**
		PSTM	.255	.229	$> .05$	5.2%
1	CLPS	ASTM SSRT	.025	.022	$> .05$	0.05%
		AC SC	.479	**.459**	$= .001$	**21.1%**
2	CLAO	ASTM SSRT	-.305	**-.302**	$< .05$	**9.1%**
		LoS	-.311	**-.309**	$< .05$	**9.5%**
		PSTM	.148	.131	$> .05$	1.7%
3	% DIS	ASTM SSRT	.193	.173	$> .05$	3.0%
		AC ErR	.246	.238	$> .05$	5.7%

Note. CLPS = perceived cross-language phonetic similarity; CLAO = cross-language assimilation overlap; % DIS = L2 vowel discrimination accuracy; ASTM SSRT = acoustic short-term memory measured by the serial sound recognition task; PSTM = phonological short-term memory; AC SC = attention control shift cost; AC ErR = attention control error rate; LoS = length of stay abroad.

short-term memory and acoustic short-term memory was partialled out. The contribution of phonological short-term memory ($beta = .255$, $Part r = .229$, $p > .05$) and acoustic short-term memory ($beta = .025$, $Part r = .022$, $p > .05$) did not reach significance level. These results suggest that taken together all three cognitive abilities significantly contribute to L2 vowel perception. Attention control appears to be the best predictor of L2 learners' ability to perceive a cross-language phonetic distance between L2 and L1 vowel sounds thus facilitating formation of new L2 phonetic categories for L2 sounds.

5.5.1.2 Cognitive abilities and perceptual assimilation overlap

As reported earlier in the chapter (see Section 5.4.1), the perceptual assimilation overlap scores were found to be related to the acoustic memory serial sound recognition scores and participants' length of stay abroad. In order to estimate the unique contribution of acoustic short-term memory to the degree of perceptual assimilation overlap between L2 vowels, a standard multiple regression analysis was carried out[7]. The regression model, which included the measures of acoustic memory and length of stay abroad as predictor

[7]Despite the assimilation overlap scores were non-normally distributed the analysis of the residuals indicated no violations of normality, linearity, and homoscedasticity and no outliers were evident(Tabachnick & Fidell, 2013).

variables and the perceptual assimilation overlap scores as a dependent variable (see Regression model 2 in Table 18), was found to be statistically significant ($F(2,42) = 5.72$, $p < .01$) and accounted for 21.4% of the variance in the degree of perceptual assimilation overlap ($R^2 = .214$, $AdjustedR^2 = .177$). When the variance explained by the length of stay abroad was partialled out ($beta = -.311$, $Partr = -.309$, $p < .05$), acoustic memory, as measured by the serial sound recognition task was found to significantly contribute ($beta = -.305$, $Partr = -.302$, $p < .05$) to the degree of phonetic distance between L2 vowels by explaining a unique 9.1% of variance in assimilation overlap scores (see Table 19). The results obtained suggest that a larger acoustic short-term memory capacity together with a longer stay in an L2-speaking country may facilitate the establishment of distinct phonetic categories for contrasting L2 sounds.

5.5.2 Contribution of cognitive abilities to L2 vowel discrimination

In order to answer the second research question, the contribution of cognitive skills to L2 vowel discrimination was estimated by performing a standard multiple regression test between L2 learners' phonological short-term memory, acoustic short-term memory serial sound recognition, and attention control error rate scores as predictor variables and discrimination accuracy scores as a dependent variable (see Regression model 3 in Table 18). As Table 18 shows, the model accounted for a significant 17.8% of variance in accurate L2 vowel discrimination ($R^2 = .178$, $AdjustedR^2 = .118$, $F(3,41) = 2.95$, $p < .05$). However, the evaluation of each one of the cognitive variables in the model revealed that neither phonological short-term memory ($beta = .148$, $Partr = .131$, $p > .05$) nor acoustic short-term memory ($beta = .193$, $Partr = .173$, $p > .05$) or attention control error rate ($beta = .246$, $Partr = .238$, $p > .05$) made a significant unique contribution to predicting accuracy in L2 vowel discrimination (see Table 19). Only attention control error rate scores were found to be closely associated with the discrimination reaction time scores ($r = -.317$, $p < .05$). The results obtained indicate that the participants who were better able to accurately shift attention between speech dimensions discriminated L2 vowels much more slowly than those who were less accurate in the attention-shifting task. Attention control, as measured by error rate, accounted for 10% of variance in predicting learners' speed in L2 vowel discrimination ($R^2 = .100$). These results suggest that accurately bringing relevant acoustic cues (i.e. spectral cues) to the perceptual foreground is a cognitively demanding task and may lead to a slower speed in the discrimination of contrasting L2 vowels.

Overall, the results obtained suggest that L2 learners' accuracy in perception of L2 vowel contrasts is predicted by a joint contribution of phonological short-term memory, acoustic short-term memory and attention control abilities. It appears that larger phonological short-term memory and acoustic short-term memory capacity, and lower attention control, operationalized as the ability to accurately bring irrelevant acoustic cues to the perceptual background, are associated with a more accurate L2 vowel discrimination. Moreover, lower ability to inhibit irrelevant acoustic cues seems to also predict learners' speed in L2 vowel discrimination.

5.6 Summary

On the whole, our results indicate that learners' perception of L2 sounds is associated with their cognitive abilities rather than with L2 proficiency and other L2 speech learning-related factors. Phonological short-term memory, acoustic short-term memory and attention control significantly contribute to L2 learners' speech perception. These cognitive abilities were found to be related to the degree of perceived phonetic distance between L2 and L1 sounds, and vowel discrimination ability. Contrary to our predictions phonological and acoustic short-term memory were associated with lower degree of perceived distance between L2 and perceptually closest L1 vowels. More efficient attention control as reflected by lower shift cost was associated with higher degree of perceived cross-language phonetic distance. Taken together, phonological short-term memory, acoustic short-term memory and attention control accounted for 33.1% of variance in the degree of perceived cross-language phonetic distance. Only attention control ability was found to make a significant unique contribution of 22.1% to the perceived distance between L2 and L1 sounds.

L2 learners' acoustic short-term memory capacity and length of stay abroad were found to be associated with lower degree of perceptual overlap between contrasting L2 sounds. The short-term memory for acoustic details and experience of stay abroad as measured in the present study were found to make a significant contribution of 9.1% and 9.5%, respectively, to the learners' higher degree of phonetic distance between contrasting L2 sounds.

Phonological short-term memory, acoustic short-term memory and attention control were also found to be related to the discrimination of L2 vowels. Larger phonological and acoustic short-term memory capacities were associated with more accurate L2 vowel perception. Attention control error rate, reflecting L2 learners' ability to accurately foreground relevant and background irrelevant acoustic cues, was found to be related to slower and less accurate L2 vowel discrimination. Only taken together phonological short-term memory, acoustic short-term memory and attention control abilities made a significant contribution of 18% to the accuracy of L2 vowel discrimination. Learners' speed in L2 vowel discrimination was explained by 10% contribution of L2 learners' lower ability to bring the relevant acoustic cues to the perceptual foreground.

On the whole, the results suggest that the cognitive abilities examined play an important role in L2 learners' ability to establish new L2 phonetic categories. These cognitive abilities seem to contribute to different extents to L2 learners' speech perception. The following chapter interprets and discusses the results obtained in light of previous research.

Chapter 6

Discussion

This dissertation investigated the role of cognitive ability in learners' acquisition of L2 sounds. Current theory in L2 acquisition along with considerable empirical evidence suggest that learners' success in acquiring L2 phonological competence is reflected in their ability to form phonetic categories for L2 sounds. In order to establish an accurate L2 sound system, learners must learn to discern the acoustic-phonetic differences between L2 and L1 sounds, and between contrasting L2 sounds. As discussed in Chapter 1, an early age of onset of L2 learning and extensive amount of L2 experience cannot explain all of the variance in L2 phonological competence.

Specifically, in this dissertation, we examined the extent to which phonological short-term memory, acoustic short-term memory and attention control may contribute to learners' ability to perceive cross-language phonetic differences between L2 and L1 sounds and their acquisition of L2 perceptual phonological competence. We hypothesized that L2 learners' individual differences in L2 speech perception may be partly attributed to the individual differences in phonological short-term memory, acoustic short-term memory and attention control. We predicted that these cognitive abilities might significantly contribute to L2 learners' ability to perceive a cross-language phonetic distance between L2 and L1 sounds, and to perceive L2 phonological contrasts.

In order to test this hypothesis, we conducted a correlational study with a group of adult Catalan-Spanish learners of English as a Foreign Language. The participants were asked to perform L2 vowel perception tasks assessing their competence in perception of L2 sounds. The participants also took part in a battery of cognitive tasks assessing their phonological short-term memory capacity, acoustic short-term memory capacity and attention control. The cognitive task scores were then related to the L2 vowel perception scores. When examining the role of individual differences in L2 speech perception we also analyzed L2 learning-related factors such as L2 proficiency, age of onset of L2 learning, amount of L2 use and length of residence in an L2-speaking environment. Overall, the results revealed that learners' cognitive skills are related to their ability to establish new phonetic categories for L2 sounds. The purpose of this chapter is to evaluate and discuss the nature of the relationship between cognitive ability and L2 speech perception in light of previous research findings. The chapter also aims to point out the limitations and the implications of this study for research and practice.

6.1 Individual differences in L2 speech perception

In this dissertation learners' L2 perceptual phonological competence was understood as the extent to which L2 learners' have established phonetic categories for L2 sounds. This is determined by their ability to distinguish between L2 and L1 sounds and indicated by their ability to discriminate contrasting L2 sounds. The results obtained showed that L2 learners do experience difficulties in learning L2 sound contrasts that do not occur in their native language. The observed individual variation in learners' L2 perceptual phonological competence scores and the comparison of the learners' performance with that of the native speakers suggests that some participants in this study might have been more successful in L2 speech learning than others and might have achieved a higher level of perceptual phonological competence than others. This section discusses the results concerning L2 learners' performance in L2 vowel perception tasks in light of previous research findings.

The results regarding the degree of perceived phonetic distance between L2 and L1 sounds showed that L2 learners readily assimilated the target English vowels /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ to the perceptually closest L1 phonetic categories /i/, /e/, /ɛ/ and /a/, and perceived most L2 vowels as instances of the closest L1 category with various degrees of goodness of fit. The observed differences in fit indexes and the degree of perceptual assimilation overlap suggest that learners perceived some L2 sounds as better fitting L1 categories than others. In general, the perceptual assimilation results support the view of current L2 speech learning models that learners' perception of L2 sounds is highly influenced by the differences between L1 and L2 phonological inventories and the perceived phonetic similarity between L1 and L2 sounds (Best & Tyler, 2007; Flege, 1995; Flege et al., 2003; van Leussen & Escudero, 2015). As predicted by the SLM and the PAM-L2, L2 learners' ability to discern differences between L2 and L1 sounds, and between contrasting L2 sounds is crucial for new L2 phonetic categories to be established. The observed inter-learner variation in the degree of perceived phonetic distance between L2 and L1 sounds and the amount of perceptual assimilation overlap between L2 sounds appears to indicate that L2 learners vary in the ability to establish phonetic categories for L2 sounds.

As for the L2 learners' ability to perceive L2 phonological contrasts, the findings of the present study are in line with previous research results showing that L2 learners have difficulties acquiring some contrasting L2 sounds (Goto, 1971; Mora & Fullana, 2007; Pallier et al., 1997). The overall L2 learners' performance on the categorical vowel discrimination task, which targeted L2 learners' perception of English /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ vowel contrasts, was significantly less accurate and slower than that of the native speakers. These results are consistent with the predictions made by previous research that contrasting L2 sounds that are mapped onto the same L1 phonetic category will be difficult to discriminate (Best & Tyler, 2007; Guion et al., 2000; Levy, 2009b). As a result, this will hinder the establishment of phonetic categories for these L2 sounds. The task used to assess L2 learners' perception of L2 phonological contrasts was designed to tap the phonological level of speech perception, which would measure the extent to which L2 learners have established new phonetic categories for the target L2 sounds (Darcy et al., 2015; Gottfried, 1984). Therefore, the L2 learners' poorer performance in this task may be explained by their inability to map the contrasting L2 sounds onto distinct L2 categories. The results also yielded considerable inter-learner variation in overall vowel discrimination accuracy and speed. There were L2 learners who performed the task at a near-native level, which suggests that there are L2 learners who are more successful in establishing new phonetic categories for L2 sounds than others.

Although the relationship between perceived cross-language phonetic distance and the perception of L2 sound contrasts was not the primary goal of this dissertation, the relationship observed between L2 vowel perception measures is worth mentioning. The two-tailed correlation analyses yielded no relationship between the degree of perceived cross-language phonetic distance, the degree of perceptual overlap between contrasting L2 sounds and L2 vowel discrimination ability. This suggests that these measures might have assessed different abilities involved in L2 speech perception. Despite being crucial to predicting and explaining L2 learners' difficulties in acquiring certain L2 sound contrasts (Guion et al., 2000; Levy, 2009b), the degree of perceived phonetic distance between L2 and L1 sounds as measured in the present study may be relatively independent from learners' perception of contrasting L2 sounds. On the other hand, the one-tailed correlation analyses revealed a significant relationship between the degree of perceptual overlap and L2 vowel discrimination accuracy. These results may mean that the extent to which perceptually similar L2 sounds overlap in the L2 learners' phonological space predicts L2 learners' ability to accurately perceive L2 phonological contrasts. This is supported by previous research suggesting that the likelihood of acquiring perceptually similar L2 sounds depends on the learners' ability to detect phonetic differences between contrasting L2 sounds (Best & Tyler, 2007).

6.2 The role of age, L2 exposure and proficiency

In order to provide a deeper analysis of individual differences in L2 perceptual phonological competence, we obtained measures of such L2 learning-related factors as age of onset of L2 learning, L2 use, length of stay abroad and L2 proficiency. In the present study neither the age of onset of L2 learning, nor amount of L2 use or L2 proficiency were found to be related to the measures of L2 vowel perception. The results obtained coincide with some previous research findings showing that early age of onset of L2 learning, large vocabulary and exposure to high-quality L2 input do not always explain inter-learner variation in L2 perception and production (Bungaard-Nielsen et al., 2011; Cerviño-Povedano & Mora, 2015; Darcy et al., 2014; Flege et al., 2006; Flege & MacKay, 2004; Højen & Flege, 2006; Pallier et al., 1997; Sebastián-Gallés & Soto-Faraco, 1999). This section aims to discuss main findings regarding the relationship between L2 learning-related factors, the perception of cross-language phonetic distance and L2 perceptual phonological competence.

Specifically, except for the L2 learners' length of stay abroad none of L2 learning-related factors were found to be related to the degree of perceptual overlap between contrasting L2 vowels. Length of stay in an English-speaking country was found to make a significant 9.5% contribution to L2 learners' degree of perceptual overlap between contrasting L2 sounds. The results suggest that a longer stay may promote L2 learners' better ability to establish distinct phonetic categories for contrasting L2 sounds perceptually equated with the same L1 category. This is in line with previous research findings showing the contribution of longer residence abroad to L2 learners' acquisition of L2 perceptual and productive competence (Flege et al., 1997; Guion et al., 2000; Purcell & Suter, 1980).

The observed lack of relationship between L2 phonological competence and L2 learning-related factors may be explained by the nature of the participants' context of L2 learning. The participants in our study were instructed learners in an EFL context. In contrast to the naturalistic setting, the foreign language classroom context does not usually provide learners with a substantial amount of high-quality L2 input which is said to play an important role in acquiring L2 pronunciation skills in both contexts (Piske, 2007). The limited amount of high-quality L2 input necessary for the development of accurate L2 phonetic categories may cause the differences in foreign language learners' age of onset of L2 learning, amount of L2 use and L2 proficiency play a less important role in L2 phonetic category formation. This idea is supported by a recent Darcy et al.'s (2016) study on L2 speech perception and production and cognitive ability in foreign language learners. The results of their study showed a significant relationship between learners' inhibitory control skill and perception and production of L2 segments. Similar to our study they also failed to observe any relationship between the demographic variables and L2 speech perception. Taken together, the results of the present and the Darcy et al.'s (in press) study may be interpreted to mean that in the foreign language learning setting, characterized by a limited amount of L2 input and a large amount of L1 exposure and use, learners' cognitive ability may play a more important role in the development of accurate L2 sound representations than factors such as age of onset, L2 use and proficiency. These L2 learning-related factors have been often found to explain individual differences in L2 phonological acquisition among immersion L2 learners. It may be the case that they play a bigger role in L2 phonological development in the naturalistic setting, where an earlier onset of exposure to a substantial high-quality L2 input and a larger amount of natural L2 use make a big difference in the development of L2 phonology, than in the instructed foreign language learning context. In other words, foreign language learners may need to use their cognitive ability in order to process a limited amount of L2 input to form accurate L2 sound representations to a larger extent than immersion L2 learners acquiring their L2 in a naturalistic context.

Another explanation for the results obtained may be the failure of the measures used in the present study to uncover the relationship between L2 perceptual phonological competence and L2 learning-related factors. Despite being used in previous L2 speech learning research as a reliable measure of L2 proficiency, L2 vocabulary size has been found to be unrelated to learners' perception of L2 vowel contrasts (Darcy et al., 2014; Cerviño-Povedano & Mora, 2015; Mora & Safronova, submitted; Safronova & Mora, 2013). The explanation for the lack of the relationship between L2 learning-related factors and the acquisition of L2 contrasts may lie in the nature of the sound contrasts we examined. Support for this idea comes from the study by Bungaard-Nielsen et al. (2011) who showed that the increasing L2 vocabulary development and extended L2 exposure may not promote acquisition of L2 vowel contrasts which are assimilated according to Single-Category and Two-Category assimilation types (see Chapter 2). As Bungaard-Nielsen et al. (2011) explain, the former is predicted to be difficult for learning and the latter is believed to be well differentiated even by the least experienced learners.

This explanation is supported by the results of the present study which yielded significant differences in accuracy and reaction times between target L2 vowel contrasts. This suggests that some of the contrasts might have been perceptually more salient and easier to discriminate than others irrespective of L2 learners' experience with the L2. Moreover, as reported in Chapter 5, we found that the cognitive skills predicted the discrimination accuracy for some of the L2 contrasts and not for others. Specifically, phonological short-term memory, acoustic short-term memory (as measured by the serial sound recognition task) and attention control (as measured by error rate) contributed significantly to learners' accurate perception of the English /ɑ/-/ʌ/ contrast. The acoustic short-term memory capacity was also found to be related to the discrimination of English /æ/ and /ʌ/. The perceptual difficulty of /ɑ/-/ʌ/ and /æ/-/ʌ/ contrasts might have required participants to use their cognitive skills to discriminate these L2 vowels to a larger extent than other contrasting L2 vowels which were more often assimilated to different L1 categories and thus were perceptually easier to differentiate.

6.3 The relationship between cognitive ability and L2 speech perception

The central question asked in this dissertation is whether phonological short-term memory, acoustic short-term memory and attention control significantly contribute to L2 learners' perception of L2 sounds. We hypothesized that these cognitive skills would significantly contribute to learners' ability to distinguish between L2 and L1 sounds, and between contrasting L2 sounds. The results obtained have confirmed our hypothesis by showing a significant relationship between phonological short-term memory, acoustic short-term memory, attention control and perception of L2 sounds. This section evaluates and discusses the contribution of each cognitive ability to learners' perception of cross-language phonetic distance and their acquisition of L2 perceptual phonological competence.

6.3.1 Contribution of cognitive ability to the degree of perceived phonetic distance between L2 and L1 sounds

As discussed in Chapter 2, cross-language phonetic similarity has long been found to affect non-native speech perception (Best, 1995) and L2 speech learning (Best & Tyler, 2007; Flege, 1995). According to previous research, L2 learners use their L1 phonological system as a blueprint for the development of their L2 phonological system (van Leussen & Escudero, 2015). This makes L2 learners' categorize L2 sounds as instances of perceptually similar L1 sounds, which hinders the establishment of accurate L2 phonetic categories. Current L2 speech learning models (see Chapter 2) suggest that learners' success in the acquisition of L2 sounds, that is, the establishment of new phonetic categories for L2 sounds is determined by their ability to perceive acoustic-phonetic differences between L2 and L1 sounds, and between contrasting L2 sounds (Best & Tyler, 2007; Flege, 1995). This dissertation is the first attempt to examine the contribution of L2 learners' cognitive skills to this ability.

This issue was addressed by our first research question: *To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' degree of perceived phonetic distance between L2 and L1 vowels?* The results obtained showed that phonological short-term memory, acoustic short-term memory and attention control (shift cost) accounted for 33.1% of variance in the L2 learners' phonetic distance scores. Only attention control ability was found to make a significant unique 22.1% contribution to learners' degree of perceived phonetic distance between L2 and L1 sounds. As expected, more efficient attention control was associated with a higher degree of perceived cross-language phonetic distance. This finding suggests that the ability to shift attention efficiently back and forth between the acoustic cues may promote the establishment of L2 phonetic categories by facilitating the detection of acoustic-phonetic differences between L2 and L1 sounds. This idea is supported by previous research findings showing that paying attention to relevant acoustic cues distinguishing sounds facilitates L2 speech learning (Darcy et al., 2014; Francis & Nusbaum, 2002; Francis et al., 2000; Guion & Pederson, 2007).

Contrary to our hypothesis both phonological short-term memory and acoustic short-term memory were found to be significantly related to a lower degree of perceived phonetic distance between L2 and L1 sounds. Our finding suggests that the short-term memory mechanisms for maintenance of phonological and acoustic information may contribute to a larger degree of perceived cross-language phonetic similarity between L2 and L1 sounds. Since this study has been the first attempt to investigate the relationship between cross-language vowel perception and cognitive ability it is difficult to explain these results in the light of previous research on L2 speech perception and cognitive ability.

One possible explanation of our findings concerns the nature of categorical perception (Lakoff, 1987; Rosch, 1978). As mentioned in Chapter 2, categorization allows the rapid and efficient information processing due to the minimization of differences between items assigned to the same category. Phonological short-term memory and acoustic short-term memory may thus aid in the categorization of L2 sounds as instances of L1 categories and make L2 learners perceive a lower degree of phonetic distance between L2 and L1 sounds to which the former are perceptually assimilated. This can be explained by the nature of working memory to hold two types of information, continuous and categorical, where the latter is critical for high working memory performance under the conditions of high memory load (Joseph et al., 2015; Li et al., 2013). Specifically, the tasks measuring the L2 learners' phonological and acoustic short-term memory were designed to place high demands on the short-term memory capacity, which might have forced participants to process speech and speech-like stimuli in a categorical manner. In the L2 vowel perceptual assimilation task requiring participants to categorize L2 vowels as instances of L1 categories and to judge the similarity between them. The participants might have used their short-term memory to keep the information that allowed them to rapidly categorize L2 sounds as instances of L1 phonetic categories.

However, due to the correlations being relatively weak, neither phonological short-term memory nor acoustic short-term memory made a significant unique contribution to L2 learners' ability to perceive phonetic distance between L2 and L1 sounds. Therefore, it appears that attentional mechanisms such as attention-shifting ability may play a bigger role in the perception of acoustic-phonetic differences between L2 and L1 sounds than the capacity to temporarily hold acoustic and phonological information in short-term memory.

Acoustic short-term memory was the only cognitive ability found to be related to L2 learners' degree of perceptual assimilation overlap between contrasting L2 sounds. Consistent with our hypothesis, the acoustic short-term memory capacity significantly contributed (9.1 %) to the degree of the perceptual assimilation overlap between contrasting L2 sounds. This finding suggests that the acoustic short-term memory capacity may play a role in the establishment of distinct phonetic categories for perceptually similar L2 sounds. Our results support previous research outcomes showing that short-term memory for acoustic properties of speech facilitates accurate L2 speech perception and production by allowing L2 learners' perceptual system to better encode acoustic details signaling differences between L2 sounds (Mora & Safronova, submitted; Safronova & Mora, 2012b; Tanaka & Nakamura, 2004).

On the whole, the findings suggest that phonological short-term memory, acoustic short-term memory and attention control are related to learners' perception of L2 sounds. Our results are consistent with research showing that short-term working memory, which comprises short-term memory storage and attentional mechanisms, plays a role in acquiring L2 speech perception and production skills (Darcy et al., 2015; Rota & Reiterer, 2009). The findings of the present study suggest that the ability to establish phonetic categories for L2 sounds is predicted by a larger acoustic short-term memory capacity and a more efficient attention control. These abilities may contribute to a more accurate L2 speech perception by promoting the detection of acoustic-phonetic properties of L2 sounds.

6.3.2 Contribution of cognitive ability to the perception of L2 phonological contrasts

The second research question addressed in this dissertation is: *To what extent do phonological short-term memory, acoustic short-term memory and attention control contribute to L2 learners' perception of L2 vowel contrasts?* We predicted that these abilities would significantly contribute to L2 learners' perception of L2 vowel contrasts. We found that phonological short-term memory, acoustic short-term memory and attention control, as measured by error rate, made a significant 18% contribution to the accuracy in perception of L2 phonological contrasts. None of the cognitive abilities were found to make a unique contribution to the perception of L2 vowel contrasts. Attention control, as measured by error rate scores, was correlated with the discrimination reaction time scores and made

10% unique contribution to explaining speed in the discrimination of L2 phonological contrasts.

As discussed in Chapter 3, a considerable body of research has demonstrated that phonological short-term memory plays an important role in L1 and L2 acquisition. In research on L2 speech learning it has also been found to significantly contribute to L2 learners' ability to establish new L2 phonetic categories (Cerviño-Povedano & Mora, 2011, 2015; MacKay et al., 2001). Results obtained in this study support previous research by demonstrating a significant correlation between phonological short-term memory capacity and the discrimination of L2 vowel contrasts. Similarly, larger acoustic short-term memory capacity was found to be associated with a more accurate perception of L2 phonological contrasts. This finding is consistent with previous research on acoustic memory showing that the ability to temporarily maintain larger amounts of acoustic information in short-term memory facilitates accurate perception and production of L2 speech (Mora & Safronova, submitted; Safronova & Mora, 2012b; Tanaka & Nakamura, 2004).

These findings fit well previous research on memory and speech perception suggesting that the discrimination of sounds involves two levels: acoustic and phonological (Fujisaki & Kawashima, 1970; Pisoni, 1973; Tanaka & Nakamura, 2004). The acquisition of L2 phonetic categories may be enhanced by the capacity to hold acoustic and phonological information in memory by means of a rehearsal mechanism. Greater memory capacity allows for the processing of a larger amount of L2 speech input to be analyzed and the acoustic-phonetic patterns between sounds to be noticed. On the basis of the detected acoustic-phonetic features new and accurate phonetic categories for L2 sounds can be inferred (Darcy et al., 2015; Skehan, 2012).

We also predicted that efficient (fast and accurate) attention-shifting skill might be related to L2 learners' more accurate perception of L2 phonological contrasts. This hypothesis was not clearly upheld. We failed to observe a relationship between learners' perception of L2 phonological contrasts and attention control as measured by the shift cost. Our results are in line with previous studies investigating L2 learners' perception of English vowel contrasts showing that smaller attention shift cost (i.e. better attention control ability) may not contribute to accurate perception of L2 phonological contrasts (Darcy et al., 2014, 2015; Safronova & Mora, 2012b). This may suggest that the ability to rapidly re-allocate focus of attention may not be involved in L2 learners' accurate perception of English vowel contrasts. Rather, the ability to focus attention on particular acoustic cues, which may be perceptually more salient than others, may be a better predictor of the L2 learners' fast and accurate discrimination of L2 vowel contrasts.

This idea is supported by our results that showed a significant relationship between L2 learners' vowel discrimination ability and attention control as measured by the error rates. Specifically, we found that L2 learners with lower attention control skill to accurately shift attention between two speech dimensions (i.e. those who had higher error rates) could discriminate L2 phonological contrasts more accurately and rapidly. Moreover, the accuracy in foregrounding relevant information explained 10% in L2 learners' speed in L2 vowel discrimination. These results suggest that learners' L2 vowel discrimination ability might be partly predicted by their lower ability to inhibit irrelevant acoustic cues. The fact that attention control accuracy and shift cost measures were found to be unrelated suggests that they measured two different constructs. That is, shift costs might have assessed attention control as a ballistic process (Segalowitz, 2010) and error rates assessed the attention-shifting dimension that involves focusing and inhibiting attention (Miyake et al., 2000).

One possible explanation of these results is that the participants who were faster and more accurate in discriminating L2 phonological contrasts might have over-relied on temporal cues when perceiving contrasting L2 vowels. This idea is in line with Bohn's (1995) Desensitization hypothesis predicting that L2 learners' may attend to duration when perceiving L2 speech contrasts whether or not this acoustic cue is employed phonologically in their L1. This has also been demonstrated by several studies with Catalan-Spanish bilingual learners of English who were found to ignore spectral differences between L2 vowels and over-rely on duration when perceiving L2 vowel contrasts (Cebrian, 2006; Cerviño-Povedano & Mora, 2011; Safronova & Mora, 2013). The fact that temporal cues are more salient and easy-to-attend than spectral cues is supported by the results of the present study. Our results showed that the two task sets in the attention-shifting task (i.e. voice quality and duration) differed in difficulty. Specifically, the participants responded much faster when shifting attention to voice quality trials than to duration trials. As explained in previous attention-shifting studies, such asymmetrical shifting costs indicate that it is easier to shift to the harder of the two tasks because larger inhibition is required to suppress the easier of the two tasks (see Costa & Santesteban, 2004; Monsell, Yeung, & Azuma, 2000, for discussion of asymmetrical shifting costs). Therefore, when distinguishing between contrasting L2 vowels, it may be perceptually easier to bring the temporal differences to the perceptual foreground than the spectral ones.

Taken together with previous research outcomes, our findings indicate that the joint contribution of phonological short-term memory, acoustic short-term memory and attention control may facilitate learners' establishment of new phonetic categories for L2 sounds (at least for L2 vowel sounds) on the basis of perceptually more salient acoustic cues, which may not be the primary cues used by native speakers. On the other hand, it may be the case that the participants in our study used their cognitive abilities to attend to the more salient temporal differences between the L2 vowels in order to discriminate the test L2 vowel contrasts in the categorical L2 vowel discrimination task. This interpretation of the results raises a question whether the ABX categorical discrimination task used in the present study measures L2 phonological competence in L2 vowel perception. Future research may need to employ various measures of L2 speech perception (Flege, 2003) to uncover the relationship between cognitive skills and learners' competence in L2 vowel perception.

Overall, the results of the present study highlight the significant contribution of short-term storage and attentional mechanisms to L2 learners' perception of L2 sounds. Interestingly, the degree of perceived phonetic distance between L2 and L1 sounds and the perception of L2 phonological contrasts appear to tap certain cognitive skills. Specifically, the efficient attention-shifting skill is the best contributor to L2 learners' ability to perceive phonetic distance between L2 and L1 sounds, whereas acoustic short-term memory significantly predicts the degree of perceptual assimilation overlap between contrasting L2 sounds. As for the perception of L2 phonological contrasts, the joint contribution of phonological short-term memory, acoustic short-term memory and attention control significantly predicts L2 learners' accuracy in the discrimination of contrasting L2 sounds. The attention control plays a role in fast and accurate discrimination of L2 sounds. These results suggest that the ability to perceive cross-language phonetic distance and the ability to discriminate contrasting L2 sounds may require different sets of cognitive mechanisms.

6.4 Limitations of the present study

The present study has some specific limitations that should be taken into account when considering its findings. The first and the most important limitation concerns the participants. First of all, a relatively small number of participants makes us evaluate the significance of the results with caution. The method of selecting the Catalan dominant participants, which was one of the factors limiting the number of participants, may need further validation. Future research may take these issues into account and involve a bigger number of participants.

The second limitation is related to the design of the cognitive and L2 vowel perception tasks. The speech samples in the perceptual assimilation task were not examined by native speakers to ensure that only good examples of English and Catalan vowels were used as stimuli in the task (Guion et al., 2000). This might have had an effect on the overall judgment of vowels by the participants in the present study. The instructions given to the participants in the perceptual assimilation task might have influenced the results regarding the relationship between cognitive abilities and cross-language phonetic distance. Specifically, the participants were instructed to judge the similarity between the

L2 sounds and L1 sounds (Lengeris, 2009). As pointed out by Bohn (2002), the instructions to rate either similarity or dissimilarity may affect the results because similarity and difference are not the same. In the perceptual assimilation task used in our study the highest point of the rating scale was the label “identical”. Thus, the way the scale was set up might have forced participants to employ their phonological short-term memory and acoustic short-term memory to attend to similarities between the L2 and L1 sounds which may not be the same as perceiving distance between the sounds. Future research may solve this issue by asking participants to judge the goodness of fit between L2 speech stimuli and L1 categories by deciding if the stimuli are good or bad examples (Guion et al., 2000) or whether the stimuli are more L1 or L2 sounding (Levy, 2009a).

Despite our designing the L2 vowel discrimination task to be categorical by using a long interstimulus interval, multiple voices and disyllabic nonwords, participants might have relied on temporal cues while discriminating English tense-lax vowels. Future research can use other measures of categorical L2 speech perception such as AXB or oddity discrimination task (Flege, 2003) as well as various interstimulus intervals to ensure categorical nature of the task. For example, Højen and Flege (2006) used a shorter or 0 ms interstimulus interval. Another way future research may solve this issue is by manipulating vowel duration when testing L2 learners’ perception of English tense-lax vowel contrasts (Mora & Safronova, submitted; Moya-Galé & Mora, 2012; Cerviño-Povedano & Mora, 2011).

Phonological short-term memory was measured by using a serial nonword recognition based on recognition of CVC nonwords (Cerviño-Povedano & Mora, 2011). The participants might have focused on vowels when deciding if there was a change in order of the nonwords, which might have led to the correlation with acoustic short-term memory. Acoustic short-term memory is thought to be particularly important for vowel perception due to the fact that vowel traces may survive longer than those of consonants (Repp, 1984). Other phonological memory tasks such as a digit span (Gathercole, Hitch, & Martin, 1997) and a nonword repetition task (Gathercole et al., 1997; Masoura & Gathercole, 1999) can be used by future studies to confirm or reject the results of the present study. As regards acoustic short-term memory measured by means of rotated speech stimuli, more research is needed to further validate this new method of assessing short-term memory for speech-related acoustic details.

Last but not least, the results of the present study may be biased by using vowels as stimuli in the tasks assessing L2 perceptual phonological competence. The cognitive abilities investigated in this study might be making a differential contribution to the perception of vowels and consonants (Repp, 1984), as well as to other dimensions of L2 phonology such as word stress or phonotactics (Darcy et al., 2015). Thus the results obtained should be interpreted with caution and should not be generalized to all aspects of L2 perceptual phonological competence. Future research may address this issue by examining learners' competence in the perception of L2 consonant contrasts and suprasegmentals.

6.5 Theoretical and practical implications

Despite the limitations outlined above, the present study has several implications which may be useful for SLA research as well as L2 teaching and learning practice. In the present section we point out major theoretical and practical implications of our results.

Overall, our findings lend further support to the evidence that an individual's cognitive ability plays a role in L2 speech learning (Darcy et al., 2014, 2015; Guion & Pederson, 2007; MacKay et al., 2001; Mora & Safronova, submitted). Our results indicate that individual differences in phonological short-term memory, acoustic short-term memory and attention control are significantly related to inter-learner variation in L2 speech perception. Future research on L2 phonological acquisition might consider the results of the present study useful when accounting for L2 learners' individual differences in L2 phonological acquisition.

Our findings imply that learners' success in L2 speech learning is at least to some extent predicted by acoustic short-term memory and attention-shifting skills which may facilitate learners' ability to distinguish between L2 and L1 sounds, and contribute to lower degree of perceptual overlap between L2 sounds assimilated to the same L1 sound. L2 learners' ability to distinguish between L2 and L1 sounds and between contrasting L2 sounds which allows learners to eventually create new categories for L2 sounds constitutes one of the key tenets of current L2 speech learning models (see Chapter 2). However, the models do not account for the mechanisms that underlie this ability. Our study has highlighted the importance of cognitive ability in L2 speech perception, which may have an impact on the L2 speech learning models. Current L2 speech learning models may incorporate the idea that learning L2 sounds in the foreign language learning setting may require learners' cognitive skills such as attention shifting and acoustic short-term memory.

Another theoretical implication concerns current models of working memory. This study investigated two well-known components of working memory for speech processing, namely, phonological short-term storage and attention control (Baddeley, 2000, 2012). Following the line of previous research on memory and speech, we predicted that auditory working memory would comprise a short-term memory store for acoustic information (Joseph et al., 2015; Tanaka & Nakamura, 2004; Williamson et al., 2010). Our findings add to the existing evidence that acoustic short-term memory is implicated in speech processing (Fujisaki & Kawashima, 1970; Joseph et al., 2015) and L2 learners' speech perception (Mora & Safronova, submitted; Safronova & Mora, 2012b). Thus, our findings may add to research on working memory, its organization and role in speech learning by proposing acoustic short-term memory as a potential component of working memory for speech.

Regarding practical implications, the present study indicates that better cognitive abilities may promote L2 speech learning. This can raise L2 teachers' and L2 learners' awareness of the role of cognitive ability in L2 speech learning and inspire them to enrich L2 learning practice by including exercises on working memory. In particular, L2 pronunciation training (Bradlow et al., 1999; Hazan et al., 2005), proved to be useful for improving L2 learners' perception and production skills, may be enriched by including working memory training exercises especially those targeting phonological, acoustic memory and attention control skills. Those exercises may lead to increasing acoustic short-term memory capacity and attention-shifting skills in order to promote L2 learners' ability to accurately perceive L2 sounds and facilitate the creation of accurate L2 phonetic categories.

Recent research has provided evidence that memory and attention can be trained. There are studies on working memory training showing that performance on working memory tasks can be significantly improved by training (Mezzacappa & Buckner, 2010; Clarady et al., 2009), and may facilitate improvement of language skills. For instance, Clarady et al. (2009) used a 10 hour computer-based brain training program by Posit Science to train nine foreign language practitioners. The program consisted of combinations of different kinds of 15-minute training exercises focusing on auditory working memory and language-related memory. One type of exercise included recognition of sequences of either audio tones or phonemes that increased in length and the item similarity. After 10 hours of training the participants showed a tendency to improve their auditory working memory and language skills such as reading and listening comprehension.

Finally, although the present study did not directly examine learners' cue weighting in L2 speech perception, our results suggest that cognitive skills may affect L2 learners' ability to accurately weight acoustic cues such as spectral quality and duration when acquiring L2 phonological contrasts. Specifically, greater short-term memory for phonological and acoustic information and lower ability to background irrelevant acoustic cues is related to L2 learners' over-reliance on salient temporal cues when perceiving English tense-lax vowel contrasts. Lending further support to the evidence for EFL learners' over-reliance on duration (Cebrian, 2006; Cerviño-Povedano & Mora, 2011; Safronova & Mora, 2013), our findings may be important for research on L2 phonological acquisition and for L2 pronunciation training. In particular, when assessing subjects' L2 phonological competence, future SLA studies on EFL learners' vowel perception may take into account the possibility that L2 learners establish L2 vowel categories on the basis of perceptually more salient acoustic cues. L2 pronunciation training programs may include exercises that train L2 learners' attention skills. This may help learners direct their attention to relevant acoustic cues signaling differences between L2 sounds, which will promote the learners' acquisition of L2 phonology (Francis et al., 2000; Guion & Pederson, 2007; Kondaurova & Francis, 2010).

Chapter 7

Conclusion

In this dissertation we aimed to shed light on the role of cognitive ability in the acquisition of L2 perceptual phonological competence, which has been relatively under-researched to date. In order to do so, we assessed the contribution of phonological short-term memory, acoustic short-term memory and attention control to L2 learners' degree of perceived cross-language phonetic distance and their perception of L2 phonological contrasts. Specifically, we looked at the role of these cognitive abilities in Catalan-Spanish bilingual EFL learners' perception and acquisition of English vowels in the /i/-/ɪ/, /ɪ/-/ɛ/, /ɑ/-/ʌ/, /ɑ/-/æ/ and /ʌ/-/æ/ contrasts. To achieve our goals, we set out an experiment, in which a group of 45 adult Catalan-Spanish bilinguals, who started learning English at different ages and differed in the amount of experience with English, took part in a battery of tests assessing their L2 vowel perception and cognitive skills.

Overall, the results obtained indicate an important role of phonological short-term memory, acoustic short-term memory and attention control in L2 learners' acquisition of perceptual phonological competence. This dissertation makes several important contributions to the current knowledge about L2 speech perception and L2 phonological acquisition. This chapter concludes the present dissertation by outlining the key findings and their potential contribution to the field of SLA as well as by suggesting directions for future research.

7.1 Contribution of the dissertation

In Chapter 2 we paid special attention to the perceptual mechanisms underlying L2 speech perception and determining learners' success in learning L2 sounds. As suggested by previous research, over the course of L1 phonological development, the L1 phonological system turns into a perceptual sieve that filters out acoustic-phonetic features not relevant for L1 speech perception. This makes L2 sounds become perceptually assimilated (i.e. categorized) to existing L1 categories, which blocks the ability to accurately perceive L2 sound contrasts. This difficulty appears to persist during L2 speech learning as perceived cross-language phonetic similarities impede the establishment of new L2 phonetic categories, that is, the acquisition of L2 perceptual phonological competence. Crucially, in order to acquire L2 sounds, L2 learners must be able to distinguish between L2 and L1 sounds and learn the acoustic-phonetic properties of contrasting L2 sounds (Best & Tyler, 2007; Flege, 1995; van Leussen & Escudero, 2015. However, little is known so far about the factors underlying this ability.

The findings of this dissertation make several important contributions. The first concerns the relationship between cognitive ability and L2 learners' degree of perceived phonetic distance between L2 and L1 sounds. To our knowledge, this is the first study in the field of L2 speech research which has examined the role of L2 learners' cognitive ability in their perception of cross-language phonetic similarity. Our findings show that cognitive abilities, such as phonological short-term memory and acoustic short-term memory, are significantly related to the perceptual assimilation of L2 sounds to L1 categories, suggesting that larger phonological and acoustic memory capacities may facilitate L2 learners' perception of L2 sounds as examples of L1 sounds. On the other hand, the attention shifting appears to be a unique significant predictor of L2 learners' ability to discern phonetic differences between L2 and L1 sounds. Moreover, acoustic short-term memory significantly contributes to the phonetic distance between perceptually similar L2 sounds in L2 learners' phonological space. Therefore, the L2 learners' ability to establish new L2 phonetic categories may require attention control and short-term memory for acoustic properties of L2 sounds.

Another important contribution is related to learners' competence in the perception of L2 phonological contrasts. Our results confirm that L2 learners do have difficulties acquiring L2 phonological contrasts. We found that Catalan-Spanish EFL learners' significantly differed from native speakers of English in the ability to discriminate contrasting English vowels. The inter-learner variation in the perception of L2 vowel contrasts was significantly related to individual differences in phonological short-term memory, acoustic short-term memory and attention control. The interaction of these cognitive abilities appears to significantly predict learners' accuracy in the perception of L2 phonological contrasts. However, our findings suggest that better cognitive abilities may promote L2 learners' acquisition of a "pseudo" native-like L2 phonological competence by facilitating the establishment of L2 phonological categories on the basis of more perceptually salient acoustic cues.

The findings of this dissertation shed new light on the role of cognitive ability in the acquisition of L2 perceptual phonological competence. Despite the strong evidence reported in this dissertation, more research is needed in order to better understand the exact nature of the relationship between cognitive skills and L2 speech perception and to examine the contribution of cognitive skills to other aspects of L2 phonological competence. In the following section we lay out potential avenues for future research.

7.2 Future research avenues

While attempting to accomplish its goals this dissertation brings to light other interesting questions related to the role of individual differences in cognitive ability and L2 phonological acquisition. These questions deserve to be addressed in future research. First and foremost this dissertation has dealt with the segmental dimension of L2 phonology and L2 learners' perception of L2 vowels, in particular. Specifically, we assessed L2 learners' perception of English vowel contrasts which L2 learners of English may perceive with over-reliance on duration (Cebrian, 2006; Cerviño-Povedano & Mora, 2011; Bohn & Flege, 1990). Investigating the role of cognitive ability in L2 vowel perception of L2 learners of English may thus fail to uncover the potential role of cognitive skills such as attention control in the acquisition of L2 phonological competence (Darcy et al., 2014, 2015; Safronova & Mora, 2013). This calls for future research with L2 learners of other languages such as L1-English learners of Spanish (Darcy et al., 2014). In their study Darcy et al. (2014) found that efficient attention control was related to L1-English learners" of Spanish but not L1-Spanish learners' of English perception of L2 vowel contrasts.

Since the categorical perception of vowels and consonants might be essentially different (see Repp, 1984, for a review), the contribution of the cognitive skills investigated in this dissertation may appear to be different for vowels and consonants. Future research may address this issue by assessing L2 learners' perception of different classes of L2 sounds. Moreover, L2 phonological competence comprises learners' ability to perceive and produce segmental and suprasegmental features of L2 speech such as intonation, rhythm and stress patterns. There has been few attempts to investigate the relationship between cognitive abilities and various dimensions of L2 phonology (Darcy et al., 2015). Thus, future research may extend this study to other components of L2 perceptual phonological competence and examine the contribution of cognitive ability to L2 speech production.

In this dissertation we assessed learners' L2 perceptual phonological competence at a single point in time. Future research may also address the issue of the relationship between cognitive ability and L2 phonological competence by adopting a longitudinal approach. This may help shed more light on the contribution of learners' cognitive skills at different stages of the development of L2 phonological competence. Moreover, longitudinal studies which would involve training L2 learners' cognitive and L2 cue-weighting skills are needed. Such studies could investigate whether L2 learners may improve their cognitive skills and use them to weigh L2 acoustic cues in a target-like manner.

7.3 Concluding remarks

We opened this dissertation with an epigraph quoting Leonardo da Vinci: "All our knowledge is the offspring of our perceptions". Our findings add to the evidence that this holds true for L2 learners' phonological knowledge, which is determined by the way they perceive L2 sounds. On the whole, the results of this dissertation show that such cognitive skills such as phonological short-term memory, acoustic short-term memory and attention control are involved in learners' perception of cross-language phonetic distance between L2 and L1 sounds, and appear to play a role in their acquisition of L2 perceptual phonological competence. Still, more research is needed to further investigate the nature of this relationship. This will help advance our knowledge of the factors promoting L2 phonological acquisition.

Bibliography

Aliaga-Garcia, C., Mora, J. C., & Cerviño-Povedano, E. (2011). Phonological short-term memory and L2 speech learning in adulthood. In K. Dziubalska-Kolaczyk (Ed.), *Poznan Studies in Contemporary Linguistics* (Vol. 47, pp. 1–14).

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*(2), 233–250.

Astle, D. E., & Scerif, G. (2009). Using developmental cognitive neuroscience to study behavioral and attentional control. *Developmental Psychobiology*, *51*(2), 107–118.

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*(4), 537–552.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning and Motivation*, *2*, 89-195.

Atkinson-King, K. (1980). Glossary. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 1 Production, pp. 277–273). New York, NY: Academic Press.

Baddeley, A. D. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baddeley, A. D. (1996a). The concept of working memory. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 1–27). Hove: Psychology Press.

Baddeley, A. D. (1996b). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *49*(1), 5–28.

Baddeley, A. D. (1997). *Human memory: Theory and practice*. Hove: Psychology Press.

Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D. (2002a). Fractionating the central executive. In D. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 246–260). New York, NY: Oxford University Press.

Baddeley, A. D. (2002b). Is working memory still working? *European Psychologist*, *7*(2), 85.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Baddeley, A. D. (2007). *Working memory, thought, and action*. New York, NY: Oxford University Press.

Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158–173.

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child–adult differences in second-Language phonological learning: The role of cross-language similarity. *Language and Speech*, *51*(4), 317–342.

Beddor, P. S., & Strange, W. (1982). Cross-language study of perception of the oral–nasal distinction. *The Journal of the Acoustical Society of America*, *71*(6), 1551–1561.

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

Best, C. T. (1995). A direct realist view on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-*

language research (pp. 171–204). Timonium, MD: York Press.

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

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

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

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*. Retrieved from https://sites.la.utexas.edu/bilingual/

Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *Journal of Speech and Hearing Research*, *15*, 5–41.

Boersma, P., & Hamann, S. (2009). *Phonology in perception*. Berlin: Mouton de Gruyter.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer (Version 5.1. 26)[Computer program]*. Retrieved from http://www.fon.hum.uva.nl/praat/

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* (p. 279—304). Timonium, MD: York Press.

Bohn, O.-S. (2002). On phonetic similarity. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 191–216). Wissenschaftlicher Verlag Trier.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*(03), 303–328.

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

Bradlow, A. R., Akahane-Yamada, R., 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.

Bungaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics*, *32*, 51–67.

Carbonell, J. F., & Llisterri, J. (1999). Catalan. In *Handbook of the international phonetic association: A guide to the usage of the international phonetic alphabet* (p. 61-–65). Cambridge: Cambridge University Press.

Cebrian, J. (2002). *Phonetic similarity, syllabification and phonotactic constraints in the acquisition of a second language contrast* (Doctoral dissertation, University of Toronto). Retrieved from http://twpl.library.utoronto.ca/index.php/twpl/article/view/6505

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

Cebrian, J. (2007). Old sounds in new contrasts: L2 production of the English tense-lax vowel distinction. In J. Trouvain & B. W. J. (Eds.), *Proceedings of the 16th International Congress of the Phonetics Sciences* (pp. 1637–1640). Saarbrucken: University of Saarland.

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

Cebrian, J., Mora, J. C., & Aliaga-Garcia, C. (2011). Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks. In M. W. Katarzyna Dziubalska-Kołaczyk & M. Kul (Eds.), *Achievements and per-*

spectives in the acquisition of second language speech: New Sounds 2010 (Vol. 1, pp. 41–52). Frankfurt am Main: Peter Lang.

Cerviño-Povedano, E., & Mora, J. C. (2011). Investigating Catalan learners of English over-reliance on duration: Vowel cue weighting and phonological short-term memory. In K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (Vol. 1, pp. 53–64). Frankfurt am Main: Peter Lang.

Cerviño-Povedano, E., & Mora, J. C. (2015). Spanish EFL learners' categorization of /iː-ɪ/ and phonological short-term memory. *Procedia - Social and Behavioral Sciences*, *173*, 18–23.

Clarady, C. K., Bunting, M. F., Novick, J., Atkins, S., Dougherty, M. R., & Forsyth, B. (2009). *Working memory training for foreign language practitioners: Study findings* (Tech. Rep.). Center for Advanced Study of Language, University of Mariland.

Conboy, B. T., Sommerville, J. A., & Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, *44*(5), 1505–1512.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, *96*(2), 341–370.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Cowan, N. (2005). *Working memory capacity*. Hove: Psychology Press.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? In W. S. Sossin, J. C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Progress in brain research* (Vol. 169, pp. 323–338). Amsterdam: Elsevier.

Cowan, N., & Morse, P. A. (1986). The use of auditory and phonetic memory in vowel discrimination. *The Journal of the Acoustical Society of America*, *79*(2), 500–507.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, *5*(6), 365–373.

Darcy, I., Mora, J. C., & Daidone, D. (2014). Attention control and inhibition influence phonological development in a second language. *Concordia Working Papers in Applied Linguistics*, *5*, 115–129.

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*, *66*, 1-33. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/lang.12161/full

Darcy, I., Park, H., & Yang, C. (2011). Suppression of L1 influence in L2 phonological processing: Cognitive abilities and individual variation. Paper presented at the Second Language Research Forum, Ames, IA.

Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, *40*, 63–72.

Darwin, C. J., & Baddeley, A. D. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology*, *6*(1), 41–60.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4084861/

Diamond, A., Werker, J. F., & Lalonde, C. (1994). Toward understanding commonalities in the development of object search, detour navigation, categorization, and speech perception. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 380–426). New York, NY: Guilford Press.

Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, 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.

Dogil, G., & Reiterer, S. M. (Eds.). (2009). *Language talent and brain activity: Trends in applied linguistics*. Berlin: Walter de Gruyter.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford: Blackwell.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306.

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

Field, A. (2009). *Discovering statistics using SPSS: and sex and drugs and rock ’n’ roll* (Third ed.). London: Sage publications.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics*, *8*(2), 162–177.

Flege, J. E. (1988). The production and perception of speech sounds in a foreign languages. In H. Winitz (Ed.), *Human communication and its disorders: A review* (pp. 224–401). Norwood, NJ: Ablex.

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

Flege, J. E. (2003). A method for assessing the perception of vowels in a second language. In E. Fava & A. Mioni (Eds.), *Issues in clinical linguistics* (pp. 19–44). Unipress: Italy.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.), *Laboratory phonology* (Vol. 9, pp. 353–380). Berlin:

Mouton de Gruyter.

Flege, J. E. (2009). Give input a chance. In M. Young-Scholten & T. Piske (Eds.), *Input matters* (pp. 175–190). Clevedon: Multilingual Matters.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and dults. *Journal of Phonetics*, *34*, 153–175.

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

Flege, J. E., & Hillenbrand, J. (1987). Limits on pronunciation accuracy in foreign language speech production. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 176–203). Rowley, MA: Newbury House.

Flege, J. E., & Liu, S. (2001). The effects of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*(04), 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., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, *106*(5), 2973–2987.

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

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

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, *62*(8), 1668–1680.

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*(2), 349.

Frankish, C. (2008). Precategorical acoustic storage and the perception of speech. *Journal of Memory and Language*, *58*(3), 815–836.

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*(3), 463–487.

Friedrich, F. J. (1990). Multiple phonological representations and verbal short-term memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 74–93). Cambridge: Cambridge University Press.

Fujisaki, H., & Kawashima, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokio*, *29*, 207–214.

Fullana, N. (2006). The development of English (FL) perception and production skills: Starting age and exposure effects. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 41–64). Clevedon: Multilingual Matters.

García Lecumberri, M. L., & Gallardo, F. (2003). English FL sounds in school learners of different ages. In M. P. García Mayo & M. L. Garcia Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 115–135). Clevedon: Multilingual Matters.

Gathercole, S. E., Hitch, G. J., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*, 966–979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A.-M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*(1), 65–77.

Gathercole, S. E., & Thorn, A. S. C. (1998). Phonological short-term memory and foreign language learning. In A. F. Healy & L. E. J. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 141–158). Mahwah, NJ: Lawrence Erlbaum.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*(5), 887–898.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds L and R. *Neuropsychologia*, *9*(3), 317–323.

Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, *12*, 91–114.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711–2724.

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning* (pp. 57–77). Amsterdam: John Benjamins.

Harnad, S. (2003). Categorical perception. In *Encyclopedia of cognitive science.* Nature Publishing Group/Macmillan. Retrieved from http://www.ecs.soton.ac.uk/~harnad/Temp/catperc.html

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360–378.

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1307–1334.

Hickok, G., Buchsbaum, B. R., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673–682.

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of*

America, *119*(5), 3059–3071.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, *30*, 225–249.

Højen, A., & Flege, J. E. (2006). Early learners' discrimination of second-language vowels. *The Journal of the Acoustical Society of America.*, *119*(5), 3072–3084.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis. *Studies in second language acquisition*, *16*(1), 73–98.

Isaacs, T., & Trofimovich, P. (2011). Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech. *Applied Psycholinguistics*, *32*(1), 113–140.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *89*.

Jilka, M., Anufryk, V., Baumotte, H., Lewandowski, N., Rota, G., & Reiterer, S. (2008). Assessing individual talent in second language production and perception. In A. S. Rauber, M. A. Watkins, & B. O. Baptista (Eds.), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech* (pp. 224–239). Florianópolis, Brazil: Federal University of Santa Catarina.

Joseph, S., Iverson, P., Manohar, S., Fox, Z., Scott, S. K., & Husain, M. (2015). Precision of working memory for speech sounds. *The Quarterly Journal of Experimental Psychology*, *68*(10), 2022-2040.

Keller, C. N., A., & Saults, J. S. (1995). Can auditory memory for tone pitch be rehearsed? *Journal of Experimental Psychology*, *21*(3), 635–645.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, *30*(3), 859–873.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, *38*(4), 569–587.

Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*.

Oxford: Pergamon Press.

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. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000.

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* (pp. 121–154). Timonium, MD: York Press.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), F13–F21.

Ladefoged, P. (2012). *Vowels and consonants* (Third ed.). Malden, MA: Blackwell.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Lalonde, C. E., & Werker, J. F. (1995). Cognitive influences on cross-language speech perception in infancy. *Infant Behavior and Development*, *18*(4), 459–475.

Larson-Hall, J. (2009). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.

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*(3), 169–187.

Lev Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *Journal of Phonetics*, *47*, 36–46.

Levy, E. S. (2009a). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *The Journal of the Acoustical Society of America*, *125*(2), 1138–1152.

Levy, E. S. (2009b). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, *126*(5), 2670–2682.

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.

Li, D., Cowan, N., & Saults, J. S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, *75*(1), 145–160.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*(3), 251–285.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics*, *46*(2), 187–200.

MacKay, I. R., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*(1-2), 103–125.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum.

Masoura, E., & Gathercole, S. E. (1999). Phonological short-term memory and foreign vocabulary learning. *International Journal of Psychology*, *34*, 383-388.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

McAllister, R. (1997). Perceptual foreign accent: L2 user's comprehension ability. In A. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 119–132). Berlin: Mouton de Gruyter.

McAllister, R. (2001). Experience as a factor in L2 phonological acquisition. *Working Papers in Linguistics*, *49*, 116–119.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, *30*(2), 229–258.

Meara, P. M. (2005). *X_Lex: the Swansea Vocabulary Levels Test. v2.05.* Swansea: Lognostics. Retrieved from http://www.lognostics.co.uk/tools/index.htm

Meara, P. M., & Miralpeix, I. (2006). *Y_Lex: the Swansea Advanced Vocabulary Levels Test. v2.05.* Swansea: Lognostics. Retrieved from http://www.lognostics.co.uk/tools/index.htm

Mezzacappa, E., & Buckner, J. C. (2010). Working memory training for children with attention problems or hyperactivity: A school-based pilot study. *School Mental Health*, *2*(4), 202–208.

Miralpeix, I. (2012). *X_Lex and Y_Lex: A validation study*. Paper presented at the 22nd Vocabulary Acquisition Research Group Conference (VARG), Newtown, UK.

Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Lawrence Erlbaum.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active* *maintenance and executive control*. Cambridge: Cambridge University Press.

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340.

Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, *7*(3), 134–140.

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research*, *63*(3-4), 250–264.

Mora, J. C. (2008). Methodological issues in assessing L2 perceptual phonological competence. In *Proceedings of the PTLC 2007 Phonetics Teaching and Learning Conference* (pp. 1–5). London: Department of Phonetics and Linguistics, University College London. Retrieved from http://www.phon.ucl.ac.uk/ptlc/ptlc2007_web_procindex.html

Mora, J. C., & Fullana, N. (2007). Production and perception of English /i/-/ɪ/ and /æ/-/ʌ/ in a formal setting: Investigating the effects of experience and staring age. In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (Vol. 3, pp. 1613–1616). Saarbrücken: Universität des Saarlandes.

Mora, J. C., & Safronova, E. (submitted). Individual differences in the perception of L2 sounds and foreign accent. In C. Wright, T. Piske, & M. Young-Scholten (Eds.), *Mind matters in SLA*. Bristol: Multilingual Matters.

Moya-Galé, G., & Mora, J. C. (2012). *Enhancing L2 vowel perception: A study on vowel manipulation and neutralization learning effects*. Paper presented at the 22nd Annual Conference of the European Second Language Association (EUROSLA), Poznań, Poland.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, *21*(1), 81–108.

Moyer, A. (2009). Input as a means to a critical end: Quantity and quality in L2 phonological attainment. In M. Young-Scholten & T. Piske (Eds.), *Input matters.* Clevedon: Multilingual Matters.

Moyer, A. (2014). Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics*, *35*(4), 418–440.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, *27*(3), 377–402.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, *29*(4), 557–581.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*(3), B9–B17.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology*, *48A*(1), 98–107.

Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language learning* (pp. 301–304). Amsterdam: John Benjamins.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191–215.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, *13*(2), 253–260.

Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory & Cognition*, *3*(1), 7–18.

Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America*, *55*(2), 328–333.

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, *25*(1), 21–52.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(2), 253–262.

Prosser, S. (1995). Aspects of short-term auditory memory as revealed by a recognition task on multi-tone sequences. *Scandinavian Audiology*, *24*(4), 247–253.

Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, *30*(2), 271–287.

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

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

Ramus, F., Peperkamp, S., Christophe, A., Jacquemot, C., Kouider, S., & Dupoux, E. (2010). A psycholinguistic perspective on the acquisition of phonology. *Laboratory Phonology*, *10*(3), 311–340.

Recasens, D. (1996). *Fonètica descriptiva del català: Assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al llarg del segle XX*. Barcelona: Institut d'Estudis Catalans.

Repp, B. H. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice* (Vol. 10, pp. 243–335). New York, NY: Academic Press.

Robinson, P. (2002). Learning conditions, aptitude complexes, and SLA. In P. Robinson (Ed.), *Individual differences and instructed language learning* (Vol. 2, pp. 113–133). Amsterdam: Johns Benjamins.

Robinson, P. (2003). Attention and memory during SLA. In C. M. Doughty & M. H. Long

(Eds.), *The handbook of second language acquisition* (pp. 631–678). Oxford: Blackwells.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, *25*, 46–73.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes and aptitude test development. In M. Pawlak (Ed.), *New perspectives on individual differences in second language learning and teaching* (pp. 57–76). Oxford: Springer.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231.

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.

Rota, G., & Reiterer, S. (2009). Cognitive aspects of pronunciation talent: how empathy, mental flexibility, working memory and intelligence interact with phonetic talent. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity. Trends in applied linguistics* (pp. 67–96). Berlin: Mouton de Gruyter.

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477–1491.

Safronova, E. (2011). *Individual Differences in Cognitive Ability and L2 Vowel Perception* (Unpublished master's thesis). University of Barcelona, Spain.

Safronova, E., & Mora, J. C. (2012a). Acoustic and phonological memory in L2 vowel perception. In S. Alegre, M. Moyer, E. Pladvall, & S. Tibau (Eds.), *At a time of crisis: English and American studies in Spain* (pp. 384–390). Barcelona: Departament de Filologia Anglesa i de Germanística, Universitat Autònoma de Barcelona.

Safronova, E., & Mora, J. C. (2012b). *Acoustic memory, phonological memory and at-*

tention control in L2 speech perception. Paper presented at the Annual Conference of the American Association for Applied Linguistics (AAAL), Boston, MA.

Safronova, E., & Mora, J. C. (2013). Attention control in L2 phonological acquisition. In A. Llanes Baró, L. Astrid Ciro, L. Gallego Balsà, & R. M. Mateus Serra (Eds.), *Applied linguistics in the age of globalization* (pp. 109–114). Lleida: Universitat de Lleida.

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge: Cambridge University Press.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406.

Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. S. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *Journal of Acoustic Society of America*, *125*(3), 1737-1743.

Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech–a positron emission tomography study. *The Journal of the Acoustical Society of America*, *120*(2), 1075–1083.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, *52*(2), 240–255.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*(2), 111–123.

Segalowitz, N. (1997). Individual differences in second language acquisition. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 85–112). Mahwah, NJ: Lawrence Erlbaum.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York, NY: Routledge.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory & Cognition*, *33*(4), 644–653.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology*, *45*(1), 21–50.

Silbert, N. H., Smith, B. K., Jackson, S. R., Campbell, S. G., Hughes, M. M., & Tare, M. (2015). Non-native phonemic discrimination, phonological short term memory, and word learning. *Journal of Phonetics*, *50*, 99–119.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (Vol. 2, pp. 69–94). Amsterdam: John Benjamins.

Skehan, P. (2012). Language aptitude. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition.* London: Routledge.

Strange, W. (2002). Speech perception and language learning: Wode's developmental model of speech perception revisited. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 245–261). Trier, Germany: Wissenschaftlicher Verlag.

Strange, W. (2007). Cross-language phonetic similarity of vowels. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 35–55). Amsterdam: John Benjamins.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153–191). Amsterdam: Johns Benjamins.

Surprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *The Journal of the Acoustical Society of America*, *110*(4), 2085–2095.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (Sixth ed.). Boston, MA: Pearson Education.

Tanaka, A., & Nakamura, K. (2004). Auditory memory and proficiency of second language speaking: a latent variable analysis approach. *Psychological Reports*, *95*(3), 723–734.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*(2), 183–203.

Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology*, *9*(1), 91–96.

Trubetzkoy, N. S. (1969). *Principles of phonology*. Translated by C. A. Baltaxe. Berkeley, CA: University of California Press.

van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, *6: 1000*. Retrieved from http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01000/full#B25

Werker, J. F. (1995). Exploring developmental changes in cross-language speech perception. In L. D. Osherson D.and Gleitman & M. Liberman (Eds.), *An invitation to cognitive science, Part I: Language* (p. 87-106). Cambridge, MA: MIT Press.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*(1), 35–44.

Werker, J. F., & Pegg, J. E. (1992). Infant speech perception and phonological acquisition. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, and implications* (pp. 285–311). Timonium, MD: York Press.

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*(1), 49–63.

Wickens, C. D., Gordon, S. E., & Liu, Y. (1997). *An introduction to human factors engineering*. New York, NY: Longman.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition*, *38*(2), 163–175.

Wode, H. (1994). Nature, nurture, and age in language acquisition. *Studies in Second Language Acquisition*, *16*(03), 325–345.

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.

Appendices

Appendix A. Online bilingual language profile questionnaire (in Catalan)

The questionnaire is available online at http://goo.gl/forms/o86HS9JkdR

Questionari sobre bilingüisme CATALÀ-CASTELLÀ

Ens agradaria demanar la teva ajuda per contestar a les preguntes següents sobre el teu historial lingüístic, i sobre l'ús, actituds i competència lingüístiques en Català i Castellà. Aquesta enquesta conté 18 preguntes que pots respondre en uns 10 minuts. Això no és una prova, i per tant no hi ha respostes correctes ni incorrectes. Si et plau, contesta cada pregunta responent amb sinceritat, ja que només així podrem obtenir dades fiables d'aquesta recerca.

Moltes gràcies per la teva ajuda i participació.

Si et plau, prem el botó per començar.

*Required

I. Informació biogràfica

Nom *

Cognoms *

Data de naixement * Per exemple: 18/10/1990

Sexe *

- ○ Home
- ○ Dona

Lloc de residència actual: ciutat *

Lloc de residència actual: província *

Lloc de residència actual: país *

Lloc de naixement (ciutat) *

NO cal respondre si és el mateix que el lloc de la teva residència actual

Quant de temps has viscut / vas viure al teu lloc de naixement? *

Per exemple: 2,10 (2 anys i 10 mesos)

Quant de temps has viscut al teu lloc de residència habitual? *

Per exemple: 2,10 (2 anys i 10 mesos)

Nivell educatiu més alt assolit *

Indica'n només un

- ○ Primària
- ○ Secundària

- Batxillerat
- Formació professional - Grau Mitjà
- Formació professional - Grau Superior
- Universitat - Diplomatura, llicenciatura o grau)
- Universitat - Màster
- Universitat - Doctorat

II. Historial lingüístic

En aquesta secció et demanem que responguis a unes preguntes sobre el teu historial lingüístic. Si et plau, respon a la pregunta seleccionant la resposta apropiada al menú desplegable.

1. A quina edat i amb qui vas començar a aprendre les següents llengües?

A quina edat vas començar a aprendre CATALÀ? *

Tria una opció:

0 - Des del naixement 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

A quina edat vas començar a aprendre CASTELLÀ? *

Tria una opció:

0 - Des del naixement 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

Amb qui vas aprendre a parlar CATALÀ principalment? *

Tria una opció:

Pares, La meva parella, Familiars, Amics i/o companys de feina, Coneguts, Desconeguts

Amb qui vas aprendre a parlar CASTELLÀ principalment? *

Tria una opció:

Pares, La meva parella, Familiars, Amics i/o companys de feina, Coneguts, Desconeguts

En quina llengua va apendre a parlar la teva mare? *

- CATALÀ
- CASTELLÀ

En quina llengua va apendre a parlar el teu pare? *

- CATALÀ
- CASTELLÀ

2. Fes una estimació en % de la formació que has rebut en CATALÀ/ CASTELLÀ

Per exemple: 30 / 70 significa que has rebut el 30% de la formació en CATALÀ i el 70% en CASTELLÀ

% de CATALÀ/ CASTELLÀ durant l'educació PRIMÀRIA *

Tria NO APLICABLE si no has rebut aquest tipus de formació

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0 NO APLICABLE

% de CATALÀ/ CASTELLÀ durant l'educació SECUNDÀRIA *

Tria NO APLICABLE si no has rebut aquest tipus de formació

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0 NO APLICABLE

% de CATALÀ/ CASTELLÀ durant el BATXILLERAT *

Tria NO APLICABLE si no has rebut aquest tipus de formació

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0 NO APLICABLE

% de CATALÀ/ CASTELLÀ durant la FORMACIÓ PROFESSIONAL *

Tria NO APLICABLE si no has rebut aquest tipus de formació

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0 NO APLICABLE

% de CATALÀ/ CASTELLÀ durant la els estudis UNIVERSITARIS *

Tria NO APLICABLE si no has rebut aquest tipus de formació

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0 NO APLICABLE

3. Fes una estimació en % de l'ÚS que se'n fa del CATALÀ i el CASTELLÀ en el teu ENTORN

El teu ENTORN inclou aquells contextos comunicatius on utilitzes una llengua: a casa teva, a la feina, amb familiars... Fes una estimació del % de CATALÀ i CASTELLÀ en el teu entorn INDEPENDENTMENT de si es correspon amb teva llengua principal o no.

% d'ús de CATALÀ i CASTELLÀ a CASA (amb les persones amb les que vius) *

Per exemple: 30 / 70 significa que en aquest entorn es parla CATALÀ un 30% i CASTELLÀ un 70%

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0

% d'ús de CATALÀ i CASTELLÀ a la FEINA (amb les persones amb les que treballes) *

Per exemple: 30 / 70 significa que en aquest entorn es parla CATALÀ un 30% i CASTELLÀ un 70%

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0

% d'ús de CATALÀ i CASTELLÀ a l'ESCOLA/INSTITUT/UNIVERSITAT (amb les persones amb les que estudies) *

Per exemple: 30 / 70 significa que en aquest entorn es parla CATALÀ un 30% i CASTELLÀ un 70%

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0

% d'ús de CATALÀ i CASTELLÀ en contextos comunicatius frequents com BOTIGUES *

Per exemple: 30 / 70 significa que en aquest entorn es parla CATALÀ un 30% i CASTELLÀ un 70%

0/100 10/90 20/80 30/70 40/60 50/50 60/40 70/30 80/20 90/10 100/0

4. Quant de temps has treballat en un entorn on principalment es parla CATALÀ / CASTELLÀ?

CATALÀ? *

En ANYS

Tria una opció: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

CASTELLÀ? *

En ANYS

Tria una opció: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

III. Ús de les llengües

En aquesta secció et demanem que responguis a preguntes sobre el teu % d'ús de CATALÀ, CASTELLÀ i ALTRES LLENGÜES. Si et plau, respon a la pregunta seleccionant la resposta apropiada al menú desplegable.

5. En una setmana normal, quin percentatge del temps utilitzes les següents llengües amb els teus amics?

La suma del % d'ús de CATALÀ+CASTELLÀ ha de ser 100%.

En una setmana normal, quin percentatge del temps utilitza CATALÀ amb els teus amics? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

En una setmana normal, quin percentatge del temps utilitza CASTELLÀ amb els teus amics? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

6. En una setmana normal, quin percentatge del temps utilitzes les següents llengües amb la teva família?

La suma del % d'ús de CATALÀ+CASTELLÀ ha de ser 100%.

En una setmana normal, quin percentatge del temps utilitzes CATALÀ amb la teva família? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

En una setmana normal, quin percentatge del temps utilitzes CASTELLÀ amb la teva família? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

7. En una setmana normal, quin percentatge del temps utilitzes les següents llengües a la feina o a l'escola/institut/universitat?

La suma del % d'ús de CATALÀ+CASTELLÀ ha de ser 100%.

En una setmana normal, quin percentatge del temps utilitzes CATALÀ a la feina o a l'escola/institut/universitat? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

En una setmana normal, quin percentatge del temps utilitzes CASTELLÀ a la feina o a l'escola/institut/universitat? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. Quan parles amb tu mateix, amb quina freqüència parles amb tu mateix en les següents llengües?

La suma del % d'ús de CATALÀ+CASTELLÀ ha de ser 100%.

Quan parles amb tu mateix, amb quina freqüència parles amb tu mateix en CATALÀ? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Quan parles amb tu mateix, amb quina freqüència parles amb tu mateix en CASTELLÀ? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. Quan fas càlculs contant, amb quina freqüència contes en les següents llengües?

La suma del % d'ús de CATALÀ+CASTELLÀ ha de ser 100%.

Quan fas càlculs contant, amb quina freqüència contes en CATALÀ? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Quan fas càlculs contant, amb quina freqüència contes en CASTELLÀ? *

Tria una opció: 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

IV. Competència

En aquesta secció et demanem que evaluïs el teu nivell de competència lingüística en les següents llengües del 0 al 10. Si et plau, respon a les preguntes marcant la casella apropiada.

10. Com parles en les següents llengües?

Com parles en CATALÀ? *

Una competència "10" vol dir un nivell nadiu (sense accent). Una competència "0" vol dir que no pots parlar aquesta llengua.

	0	1	2	3	4	5	6	7	8	9	10	
no sé parlar CATALÀ	○	○	○	○	○	○	○	○	○	○	○	excel·lent

Com parles en CASTELLÀ? *

Una competència "10" vol dir un nivell nadiu (sense accent). Una competència "0" vol dir que no pots parlar aquesta llengua.

	0	1	2	3	4	5	6	7	8	9	10	
no sé parlar CASTELLÀ	○	○	○	○	○	○	○	○	○	○	○	excel·lent

11. Quin és el teu nivell de comprensió oral de les següents llengües?

"0" vol dir que no entens aquesta llengua "GENS". "10" vol dir que entens aquesta llengua "PERFECTAMENT", amb un nivell nadiu.

Quin és el teu nivell de comprensió oral en CATALÀ? *

	0	1	2	3	4	5	6	7	8	9	10	

GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

Quin és el teu nivell de comprensió oral en CASTELLÀ? *

	0	1	2	3	4	5	6	7	8	9	10	
GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

12. Com llegeixes en les següents llengües?

“0” vol dir que no pots llegir en aquesta llengua. “10” vol dir que llegeixes en aquesta llengua “PERFECTAMENT”, amb un nivell nadiu.

Com llegeixes en CATALÀ? *

	0	1	2	3	4	5	6	7	8	9	10	
GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

Com llegeixes en CASTELLÀ? *

	0	1	2	3	4	5	6	7	8	9	10	
GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

13. Com escrius en les següents llengües?

“0” vol dir que no pots escriure en aquesta llengua. “10” vol dir que escrius en aquesta llengua “PERFECTAMENT”, amb un nivell nadiu.

Com escrius en CATALÀ? *

	0	1	2	3	4	5	6	7	8	9	10	

GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

Com escrius en CASTELLÀ? *

	0	1	2	3	4	5	6	7	8	9	10	
GENS	○	○	○	○	○	○	○	○	○	○	○	EXCEL·LENT

V. Actituds

En aquesta secció et demanem que responguis a les següents afirmacions sobre actituds lingüístiques marcant una casella de l'1 al 7 Si et plau, respon a les perguntes marcant la casella apropiada.

14. Em sento "jo mateix" quan parlo en les següents llengües.

Em sento "jo mateix" quan parlo en CATALÀ. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

Em sento "jo mateix" quan parlo en CASTELLÀ. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

15. M'identifico amb les cultures següents.

M'identifico amb una cultura CATALANOPARLANT. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

M'identifico amb una cultura CASTELLANOPARLANT. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

16. És important per a mi utilitzar / arribar a utilitzar les següents llengües com un parlant nadiu.

És important per a mi utilitzar / arribar a utilitzar el CATALÀ com un parlant nadiu. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

És important per a mi utilitzar / arribar a utilitzar el CASTELLÀ com un parlant nadiu. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

17. Vull que els altres pensin que sóc un parlant nadiu de les següents llengües.

Vull que els altres pensin que sóc un parlant nadiu de CATALÀ. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

Vull que els altres pensin que sóc un parlant nadiu de CASTELLÀ. *

	1	2	3	4	5	6	7	
NO estic GENS d'acord	○	○	○	○	○	○	○	estic MOLT d'acord

18. Consentiment i anonimat

Entenc que el meu anonimat com a participant queda en tot moment garantit i que les dades que se n'obtinguin d'aquest qüestionari seràn utilitzades només per a fins científics en el marc de la recerca duta a terme per investigadors de la Universitat de Barcelona.

Consentiment *

Dono la meva autorització perquè aquestes dades puguin ser utilitzades exclusivament per a fins de recerca.

- ○ SÍ autoritzo l'ús anònim de les meves dades contingudes en aquest qüestionari
- ○ NO autoritzo l'ús anònim de les meves dades contingudes en aquest qüestionari

Appendix B. Online linguistic background questionnaire for L2 learners

The questionnaire is available at http://goo.gl/forms/d4fBVuODtg

*Required

Personal Data

Subject ID *

Please type in the CODE you were sent through e-mail here

Group *

- A
- B

Sex *

- Male
- Female

Date of Birth *

dd/mm/yyyy

You are: *

- right-handed
- left-handed
- both

Language Background

1. Indicate which language(s) you normally speak on daily basis: *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ______

2. Indicate your mother tongue (First Language or L1): *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ______

3. Indicate language(s) spoken at home most of the time: *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ______

4. Parents' First Language(s) (mother tongue or L1): *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ____________

5. Language your parents normally speak between them: *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ____________

6. Language you speak with your parents: *

You can tick more than one box

☐ Catalan

☐ Spanish

☐ English

☐ Other: ____________

7.1. Please list all the languages you know in order of dominance and age at which you started to learn them: *

For example, Catalan/0

____

7.2. *

For example, Spanish/2

____

7.3. *

For example, English/8

____

7.4.

For example, German/18

7.5.

For example, Russian/20

7.6.

For example, Greek/30

8. Please list what percentage of the time you are currently and on average exposed to each language you speak: *

Your percentages should add up to 100%. For example, Catalan/60%, Spanish/30%, English/10%

Language Use

9.1. Estimate the % of DAILY use of Catalan:*

Type in the number from 0 to 100. Please, make sure that the percentages in the next 3 questions add up to 100. For example: Catalan 70, Spanish 20 and English 10 = 100

9.2. Estimate the % of DAILY use of Spanish: *

9.3. Estimate the % of DAILY use of English:*

English Learning Experience

10. Age at which you started learning English?*

11.1. How many hours a week did you study English in Primary School? *

For example: 2 hours

11.2.How many hours a week did you study English in Secondary School? *

For example: 2 hours

11.3. How many hours a week did you study English in High School? *

For example: 2 hours

11.4. How many hours a week do you study English at University? *

For example: 2 hours

12. Have you taken English courses/private classes outside your school/university?

*

If YES, please, specify whether it was in Primary/Secondary/High school/University, where (British Council) and how long you took those classes (2 years, 4 hours a week).

13. Do you have any Certificate of English level? *

If YES, please, specify whether it is First Certificate/Advanced/Proficieny. If NO, type in NO.

14.1. Estimate the number of hours spent DAILY listening to English spoken by Non-Native speakers of English: *

For example, 2 hours

14.2. Estimate the number of hours spent DAILY listening to English spoken by Native speakers of English: *

For example, 2 hours

14.3. Estimate the number of hours spent WEEKLY listening to Non-Native speakers of English: *

For example, 14 hours

14.4. Estimate the number of hours spent WEEKLY listening to Native speakers of English: *

For example, 14 hours

14.5 Estimate the number of hours spent DAILY speaking English with Non-Native Speakers of English: *

For example, 2 hours

14.6 Estimate the number of hours spent DAILY speaking English with Native Speakers of English: *

For example, 2 hours

14.7. Estimate the number of hours spent WEEKLY speaking English with Non-Native Speakers of English: *

For example, 14 hours

14.8. Estimate the number of hours spent WEEKLY speaking English with Native Speakers of English: *

For example, 14 hours

15.1. Estimate the % of exposure to British and American English: *

For example: British 50% - American 50%

15.2. Do you think your pronunciation is more British- or American-like? *

○ British

○ American

16. For each of the items below, choose the response that corresponds to the amount of time you estimate you spend on average doing each activity in English. *

Please, use this scale: 1 - never; 2 - a few times a year; 3 - monthly; 4 - weekly; 5 - daily

	1	2	3	4	5
Watching English language television	○	○	○	○	○
Reading newspapers/magazines in English	○	○	○	○	○
Reading books in English	○	○	○	○	○
Listening to songs in English	○	○	○	○	○
Watching movies or videos in English	○	○	○	○	○
Speaking English with native or fluent speakers	○	○	○	○	○
Speaking English with nonnative speakers	○	○	○	○	○
Writing e-mails/letters in English	○	○	○	○	○

17.1. FAMILY. Please, estimate how many hours per WEEK you use English with your parents/relatives: *

17.2. WORK. Please, estimate how many hours per WEEK you use English at work: *

17.3. NATIVE FRIENDS. Please, estimate how many hours per WEEK you use English with your friends (Native Speakers of English): *

17.4. FILMS/MUSIC/BOOKS IN ENGLISH. Please, estimate how many hours per WEEK you watch/listen to/read films/music/books: *

18. Rate your command of English on the scale: 1=VERY POOR and 9=NEAR-NATIVE *

	1	2	3	4	5	6	7	8	9
Reading:	○	○	○	○	○	○	○	○	○
Listening:	○	○	○	○	○	○	○	○	○
Speaking:	○	○	○	○	○	○	○	○	○
Writing:	○	○	○	○	○	○	○	○	○
Pronunciation:	○	○	○	○	○	○	○	○	○

19. Have you ever lived in an English-speaking country more than 2 weeks? *

○ Yes

○ No

19.1. If YES, please, specify how many times, where, when, for how long and the purpose of the stay: *

If NO, please, type in ”no”

20. Have you previously received any type of Phonetic Training in English? *

- ○ Yes
- ○ No

20.1. If YES, please, specify where, when, for how long: *

If NO, please, type in ”no”

21. Do you have any speech or hearing problems/pathology? *

- ○ Yes
- ○ No

21.1. If YES, please, specify: *

If NO, please, type in ”no”

Appendix C. Linguistic background questionnaire for English native speakers

Participant's Code:	Gender:
Gender:	Place of Birth:
Age:	Residence Country/State:

1. Your mother tongue: ____________________________________

2. Estimate your proficiency level in all the language that you speak (1 = low and 10 = high):

Language	1	2	3	4	5	6	7	8	9	10
English	1	2	3	4	5	6	7	8	9	10
	1	2	3	4	5	6	7	8	9	10
	1	2	3	4	5	6	7	8	9	10
	1	2	3	4	5	6	7	8	9	10
	1	2	3	4	5	6	7	8	9	10
	1	2	3	4	5	6	7	8	9	10

3. Countries/Cities where you lived and for how long (more than 1 year):

- currently: ______________________________ since ________________________

- previously 1: ________________________ from ____________ to ____________

- previously 2: ________________________ from ____________ to ____________

- previously 3: ________________________ from ____________ to ____________

4. Languages that you are most exposed to (please, specify: 1 - little, 7 - a lot):

Language	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7

5. Parents' linguistic background

Mother's first language: ___________________ Place of Birth: ___________________

Father's first language: ___________________ Place of Birth: ___________________

Language that your parents speak with each other:_____________________________

Language that you speak with your parents: __________________________________

6. Amount of language use (%)

Estimate the % of use of English during a week (Home+...+Shops/Restaurants = 100%):

Home	100	90	80	70	60	50	40	30	20	10	0
Work	100	90	80	70	60	50	40	30	20	10	0
Relatives	100	90	80	70	60	50	40	30	20	10	0
Friends	100	90	80	70	60	50	40	30	20	10	0
Shops/Restaurants	100	90	80	70	60	50	40	30	20	10	0

Estimate the % of usage of other languages you speak (Home+...+Shops/Restaurants = 100%)

Language:______________________________________

Home	100	90	80	70	60	50	40	30	20	10	0
Work	100	90	80	70	60	50	40	30	20	10	0
Relatives	100	90	80	70	60	50	40	30	20	10	0
Friends	100	90	80	70	60	50	40	30	20	10	0
Shops/Restaurants	100	90	80	70	60	50	40	30	20	10	0

Language:______________________________________

Home	100	90	80	70	60	50	40	30	20	10	0
Work	100	90	80	70	60	50	40	30	20	10	0
Relatives	100	90	80	70	60	50	40	30	20	10	0
Friends	100	90	80	70	60	50	40	30	20	10	0
Shops/Restaurants	100	90	80	70	60	50	40	30	20	10	0

Language:______________________________________

Home	100	90	80	70	60	50	40	30	20	10	0
Work	100	90	80	70	60	50	40	30	20	10	0
Relatives	100	90	80	70	60	50	40	30	20	10	0
Friends	100	90	80	70	60	50	40	30	20	10	0
Shops/Restaurants	100	90	80	70	60	50	40	30	20	10	0

Estimate the % of usage of English:

This week	100	75-100	50-75	25-50	0-25
This month	100	75-100	50-75	25-50	0-25
This year	100	75-100	50-75	25-50	0-25
Last 5 years	100	75-100	50-75	25-50	0-25

7. Have you ever had any hearing problem or language disability?

If yes, please, explain:______________________________________

Appendix D. Linguistic background questionnaire for Catalan native speakers

Nom i Cognoms:	Data naixament:
e-mail:	Telèfon de contacte:
Lloc de naixement:	Lloc de residència habitual:

1. Llengua materna o primera: ____________________________________

2. Altres llengües que parleu (indiqueu si bé, regular, poc):

(a) ____________________________________ molt bé bé regular poc

(b) ____________________________________ molt bé bé regular poc

(c) ____________________________________ molt bé bé regular poc

3. Estimeu el vostre nivell de competència en les llengües estrangeres que parleu:

1= conec algunes paraules però no parlo mai

10= puc mantenir una conversa normal sobre temes diversos sense cap problema.

(a) ________________________ 1 2 3 4 5 6 7 8 9 10

(b) ________________________ 1 2 3 4 5 6 7 8 9 10

(c) ________________________ 1 2 3 4 5 6 7 8 9 10

4. Llocs on heu viscut i temps (superiors a un any):

actualment: ____________	des de ____________	fins a ____________
anteriorment 1: __________	des de ____________	fins a ____________
anteriorment 2: __________	des de ____________	fins a ____________

5. Llengües a les que heu estat exposats més (indiqueu si poc-1 o molt-7):

(a) ________________________ poc 1 2 3 4 5 6 7 molt

(b) ________________________ poc 1 2 3 4 5 6 7 molt

(c) ________________________ poc 1 2 3 4 5 6 7 molt

6.

Llengua materna o primera de la mare: ___________ Lloc de naixement: ___________

Llengua materna o primera del pare: ___________ Lloc de naixement: ____________

Llengua que parlen els vostres pares entre ells: _________________________

Llengua amb la que parleu amb els vostres pares: ________________________

7. Percentatge (%) d'ús habitual (marqueu una casella) de les llengües que parleu

Estimeu el % d'ús de CATALÀ en una setmana normal

Context	100	90	80	70	60	50	40	30	20	10	0
Casa	100	90	80	70	60	50	40	30	20	10	0
Feina	100	90	80	70	60	50	40	30	20	10	0
Familiars	100	90	80	70	60	50	40	30	20	10	0
Amics	100	90	80	70	60	50	40	30	20	10	0
Botigues etc	100	90	80	70	60	50	40	30	20	10	0

Estimeu el % d'ús de CASTELLÀ en una setmana normal

Context	100	90	80	70	60	50	40	30	20	10	0
Casa	100	90	80	70	60	50	40	30	20	10	0
Feina	100	90	80	70	60	50	40	30	20	10	0
Familiars	100	90	80	70	60	50	40	30	20	10	0
Amics	100	90	80	70	60	50	40	30	20	10	0
Botigues etc	100	90	80	70	60	50	40	30	20	10	0

Estimeu el % d'ús d'UNA 3ª LLENGUA ( )en una setmana normal

Context	100	90	80	70	60	50	40	30	20	10	0
Casa	100	90	80	70	60	50	40	30	20	10	0
Feina	100	90	80	70	60	50	40	30	20	10	0
Familiars	100	90	80	70	60	50	40	30	20	10	0
Amics	100	90	80	70	60	50	40	30	20	10	0
Botigues etc	100	90	80	70	60	50	40	30	20	10	0

Estimeu el % d'ús de CATALÀ habitual (encercleu l'opció) en:

una setmana normal:	100	75-100	50-75	25-50	0-25
aquest mes:	100	75-100	50-75	25-50	0-25
aquest any:	100	75-100	50-75	25-50	0-25
últims 5 anys:	100	75-100	50-75	25-50	0-25

8. Comentaris

Tens algun problema de parla i/o oïda?

NO ☐ SI ☐ Especifica:

En un futur, t'agradaria poder participar en un altre estudi?

NO ☐ SI ☐

Entenc que l'anonimat de les dades obtingudes mitjançant aquest qüestionari i aquest estudi queda en tot moment garantit i autoritzo a utilitzar aquestes dades només amb fins científics i de recerca.

_____________, ______, de ______ de 2014 _____________

Appendix E. Perceptual assimilation task materials

Table E.1

Stimuli in the perceptual assimilation task.

Stimuli	F0 (Hz)	F2 (Hz)	F1 (Hz)	Vowel duration (ms)	Stimuli	F0 (Hz)	F2 (Hz)	F1 (Hz)	Vowel duration (ms)
C1_biss1	120	2138	281	144	C1_bass3	107	1423	608	125
C1_biss2	119	2118	271	138	C1_bass4	108	1381	604	134
C1_biss3	119	2075	283	135	C2_bass1	102	1435	689	125
C1_biss4	125	2168	278	143	C2_bass2	103	1378	656	131
C2_biss1	110	2067	309	130	C2_bass3	107	1361	662	140
C2_biss2	113	2097	292	137	C2_bass4	108	1376	656	135
C2_biss3	110	2110	293	142	C3_bass1	81	1309	658	122
C2_biss4	110	2110	296	129	C3_bass2	80	1278	655	120
C3_biss1	84	2062	292	119	C3_bass3	76	1326	692	130
C3_biss2	82	2133	291	132	C3_bass4	76	1324	676	130
C3_biss3	81	2068	300	120	E1_beace1	117	2181	312	111
C3_biss4	83	2089	306	115	E1_beace2	126	2183	339	120
C1_béss1	126	1920	354	127	E1_beace3	128	2157	310	115
C1_béss2	119	1979	363	131	E1_beace4	125	2169	314	120
C1_béss3	114	2069	384	136	E2_beace1	103	2163	308	119
C1_béss4	109	2049	377	138	E2_beace2	101	2211	301	121
C2_béss1	103	1912	409	131	E2_beace3	100	2163	305	112
C2_béss2	107	1978	425	137	E2_beace4	101	2183	306	123
C2_béss3	104	1935	418	138	E3_beace1	104	2175	327	118
C2_béss4	108	1931	419	131	E3_beace2	103	2202	300	129
C3_béss1	82	1904	389	120	E3_beace3	106	2155	325	114
C3_béss2	82	1960	394	126	E3_beace4	108	2172	345	124
C3_béss3	82	1931	400	126	E1_biss1	106	1766	446	95
C3_béss4	83	1880	409	125	E1_biss2	114	1796	478	100
C1_bèss1	118	1854	480	138	E1_biss3	108	1707	473	110
C1_bèss2	119	1889	517	131	E1_biss4	115	1756	477	112
C1_bèss3	117	1802	464	142	E2_biss1	104	1804	420	106
C1_bèss4	111	1885	465	134	E2_biss2	103	1793	419	105
C2_bèss1	113	1815	519	140	E2_biss3	102	1824	405	113
C2_bèss2	106	1788	519	137	E2_biss4	106	1784	422	116
C2_bèss3	106	1829	533	131	E3_biss1	106	1795	450	100
C2_bèss4	107	1821	525	133	E3_biss2	104	1860	485	106
C3_bèss1	77	1730	547	126	E3_biss3	104	1827	454	111
C3_bèss2	79	1724	516	120	E3_biss4	107	1856	474	105
C3_bèss3	82	1755	525	125	E1_bess1	120	1592	617	113
C3_bèss4	80	1733	521	126	E1_bess2	110	1638	604	105

Stimuli	F0 (Hz)	F2 (Hz)	F1 (Hz)	Vowel duration (ms)	Stimuli	F0 (Hz)	F2 (Hz)	F1 (Hz)	Vowel duration (ms)
C1_bass1	117	1319	600	138	E1_bess3	110	1644	621	111
C1_bass2	108	1435	627	137	E1_bess4	106	1649	623	112
E2_bess1	100	1652	576	106	E1_buss3	107	1348	708	106
E2_bess2	97	1639	583	107	E1_buss4	114	1269	710	107
E2_bess3	101	1661	530	103	E2_buss1	101	1320	613	107
E2_bess4	95	1641	574	116	E2_buss2	98	1304	610	110
E3_bess1	99	1701	597	108	E2_buss3	97	1340	602	111
E3_bess2	110	1657	580	116	E2_buss4	97	1361	607	119
E3_bess3	104	1712	580	111	E3_buss1	110	1241	683	118
E3_bess4	106	1718	581	102	E3_buss2	104	1260	721	106
E1_baass1	106	1064	859	138	E3_buss3	102	1260	693	118
E1_baass2	112	1150	846	137	E3_buss4	98	1253	728	116
E1_baass3	119	1180	736	139	E1_bass1	110	1639	725	131
E1_baass4	115	1072	719	140	E1_bass2	107	1573	728	137
E2_baass1	96	1237	742	139	E1_bass3	104	1595	712	131
E2_baass2	96	1233	734	130	E1_bass4	108	1621	726	125
E2_baass3	97	1226	704	132	E2_bass1	101	1652	624	140
E2_baass4	95	1262	721	143	E2_bass2	98	1671	614	139
E3_baass1	107	1071	764	140	E2_bass3	99	1668	617	137
E3_baass2	107	1177	732	138	E2_bass4	95	1588	645	136
E3_baass3	101	1042	759	137	E3_bass1	102	1696	655	133
E3_baass4	101	1131	820	139	E3_bass2	102	1703	662	131
E1_buss1	109	1259	686	106	E3_bass3	100	1717	662	143
E1_buss2	109	1342	699	102	E3_bass4	108	1719	672	138

Note: C1, C2, C3 - Catalan stimuli, E1, E2, E3 - English stimuli

Instruccions Generals

En aquesta tasca sentiràs una sèrie de paraules inventades

La teva tasca consisteix en seleccionar la imatge que

representa la paraula que conté una vocal semblant a la que

has sentit en la paraula inventada.

Per exemple, si has sentit la paraula inventada “GAR”

i has de triar entre el dibuix d’un got i el de un gat

triaràs el dibuix del gat perquè conté la mateixa vocal que “GAR”.

Prem ESPAI per a continuar

Instruccions Generals

Llavors

has d’indicar en una escala de l’1 al 7 el grau de semblança

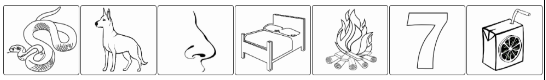
entre la vocal de la paraula inventada i

la de la paraula que has seleccionat mitjançant la imatge.

Prem ESPAI per a continuar

Instruccions pas-a-pas

Aquestes són les 7 imatges que veuràs a la pantalla



serp gos nas llit foc set suc

☐ ☐ ☐ ☐ ☐ ☐ ☐

1. Escolta la paraula (Exemple: cAs)

2. Fixa’t en la vocal (A)

3. Clica a ESCOLTA si vols escoltar la paraula

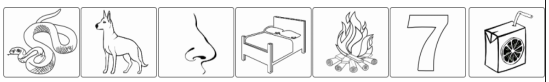
una altra vegada (només pots fer-ho una vegada)

Prem ESPAI per a continuar

Escolta

Instruccions pas-a-pas

Aquestes són les 7 imatges que veuràs a la pantalla



serp gos nas llit foc set suc

☐ ☐ ■ ☐ ☐ ☐ ☐

4. Marca la casella de sota la imatge que correspongui

Prem ESPAI per a continuar

Instruccions pas-a-pas

5. Decideix el nivell de semblança de les vocals

de la paraula inventada i la que has seleccionat clicant

sobre el nivell de l’escala que correspongui.

1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐
Diferent						**Igual**

Abans de clicar SEGÜENT assegura’t que ambdues caselles han estat seleccionades.

6. Clica a SEGÜENT per escoltar la següent paraula

SEGÜENT

Prem ESPAI per a continuar

Primer hi ha uns EXEMPLES de pràctica

RESPON TAN RÀPIDAMENT COM PUGUIS

PREPARAT/PREPARADA?

Prem ESPAI per a continuar

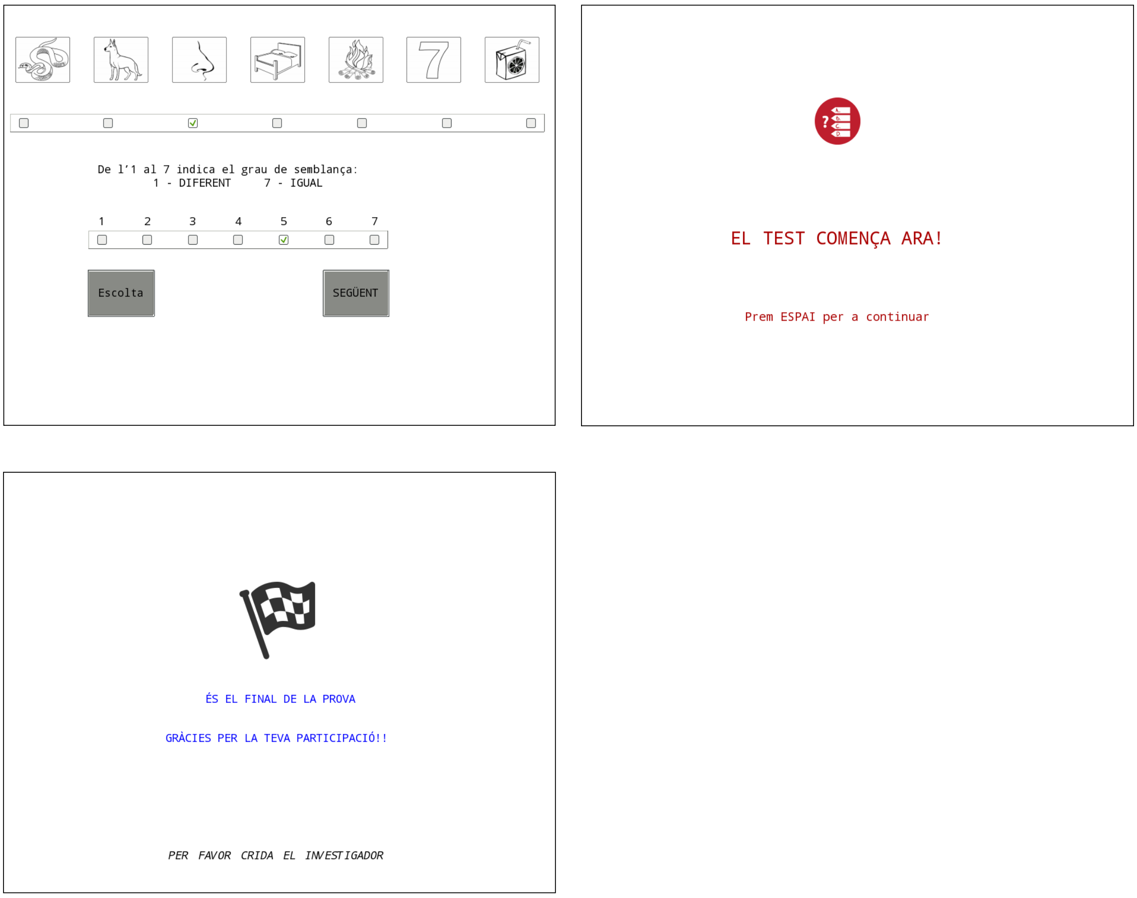


Figure E.1. Instructions in the perceptual assimilation task.

Appendix F. L2 vowel discrimination task materials

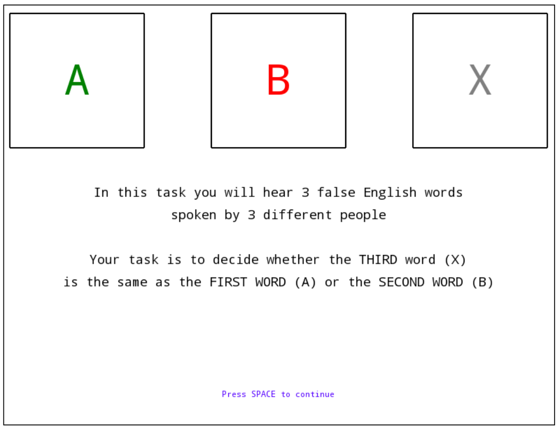
Table F.1

Stimuli in the L2 vowel discrimination task.

Stimuli	F0 (Hz)	F1 (Hz)	F2 (Hz)	Vowel duration (ms)	Stimuli	F0 (Hz)	F1 (Hz)	F2 (Hz)	Vowel duration (ms)
C1_V1_S1	156	257	2417	97	C4_V1_S1	143	238	2285	60
C1_V1_S2	159	205	2361	95	C4_V1_S2	157	215	2330	74
C1_V1_S3	130	278	2169	96	C4_V1_S3	145	273	2270	92
C1_V2_S1	150	395	2131	77	C4_V2_S1	157	432	1763	55
C1_V2_S2	154	373	2039	73	C4_V2_S2	177	477	1783	61
C1_V2_S3	173	462	1909	77	C4_V2_S3	157	406	1736	57
C1_V3_S1	141	602	1756	74	C4_V3_S1	130	616	1637	65
C1_V3_S2	134	647	1791	67	C4_V3_S2	170	647	1658	64
C1_V3_S3	149	651	1591	73	C4_V3_S3	150	649	1614	66
C1_V4_S1	122	860	1311	77	C4_V4_S1	125	772	1245	90
C1_V4_S2	132	765	1342	87	C4_V4_S2	149	703	1365	107
C1_V4_S3	150	861	1187	96	C4_V4_S3	139	872	1036	111
C1_V5_S1	119	843	1618	85	C4_V5_S1	125	763	1660	92
C1_V5_S2	152	735	1622	106	C4_V5_S2	154	721	1616	95
C1_V5_S3	151	814	1642	110	C4_V5_S3	159	760	1484	118
C1_V6_S1	146	655	1163	74	C4_V6_S1	134	602	1166	60
C1_V6_S2	168	671	1439	77	C4_V6_S2	159	620	1384	55
C1_V6_S3	156	646	1296	70	C4_V6_S3	166	762	1276	63
C2_V1_S1	145	221	2393	80	C5_V1_S1	128	256	2149	96
C2_V1_S2	153	215	2322	85	C5_V1_S2	148	200	2392	90
C2_V1_S3	129	257	2318	92	C5_V1_S3	123	216	2194	92
C2_V2_S1	157	419	1852	50	C5_V2_S1	124	414	1680	67
C2_V2_S2	163	434	1945	68	C5_V2_S2	151	381	1769	60
C2_V2_S3	149	437	1849	76	C5_V2_S3	116	431	1651	61
C2_V3_S1	141	565	1700	64	C5_V3_S1	126	616	1600	85
C2_V3_S2	149	639	1638	70	C5_V3_S2	138	612	1547	77
C2_V3_S3	149	632	1643	71	C5_V3_S3	118	628	1531	72
C2_V4_S1	117	862	1307	90	C5_V4_S1	108	754	1146	94
C2_V4_S2	145	708	1162	100	C5_V4_S2	128	719	1408	125
C2_V4_S3	149	774	1038	107	C5_V4_S3	126	853	1120	113
C2_V5_S1	115	786	1629	90	C5_V5_S1	110	772	1634	108
C2_V5_S2	142	737	1547	123	C5_V5_S2	126	755	1603	122
C2_V5_S3	147	796	1656	119	C5_V5_S3	114	838	1439	125
C2_V6_S1	142	598	1107	67	C5_V6_S1	110	593	1120	68
C2_V6_S2	171	685	1194	60	C5_V6_S2	134	646	1398	70
C2_V6_S3	166	629	1386	61	C5_V6_S3	134	735	1250	64

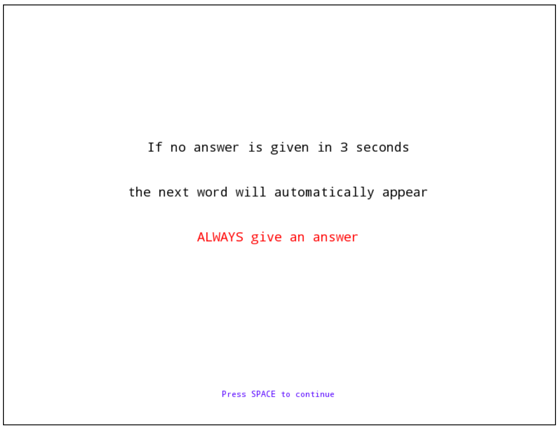
Stimuli	F0 (Hz)	F1 (Hz)	F2 (Hz)	Vowel duration (ms)	Stimuli	F0 (Hz)	F1 (Hz)	F2 (Hz)	Vowel duration (ms)
C3_V1_S1	135	221	2368	92	C6_V1_S1	135	268	2392	93
C3_V1_S2	148	237	2407	96	C6_V1_S2	130	292	2355	95
C3_V1_S3	128	268	2287	100	C6_V1_S3	131	327	2299	94
C3_V3_S3	130	670	1552	82	C6_V3_S3	122	737	1750	84
C3_V4_S1	106	854	1382	100	C6_V4_S1	104	803	1219	92
C3_V4_S2	134	761	1375	150	C6_V4_S2	112	732	1258	120
C3_V4_S3	134	702	1399	135	C6_V4_S3	134	921	1153	121
C3_V5_S1	116	769	1705	97	C6_V5_S1	110	839	1759	110
C3_V5_S2	144	691	1722	127	C6_V5_S2	111	696	1670	133
C3_V5_S3	140	784	1651	130	C6_V5_S3	118	781	1681	131
C3_V6_S1	121	668	1306	70	C6_V6_S1	110	657	1147	68
C3_V6_S2	146	652	1482	70	C6_V6_S2	136	695	1347	69
C3_V6_S3	145	704	1376	81	C6_V6_S3	134	623	1388	70

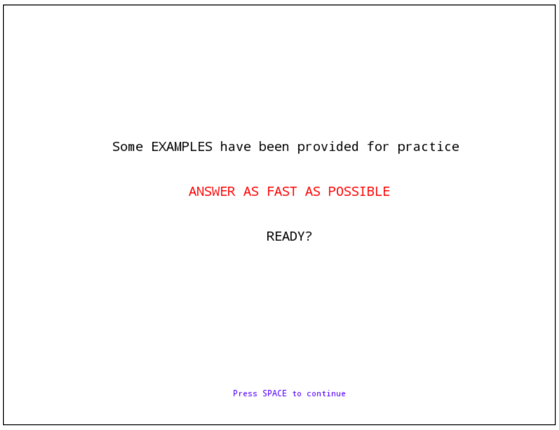
Note: C1 - [ˈp_kəs], C2 -[ˈk_fɪv], C3 - [ˈs_tʃən], C4 - [ˈt_sɪʃ], C5 - [ˈl_tɪf], C6 - [ˈm_kət], S = speaker, V1 - /i/, V2 - /ɪ/, V3 - /ɛ/, V4 - /ɑ/, V5 - /ʌ/, V6 - /æ/


A
B
X
In this task you will hear 3 false English words
spoken by 3 different people
Your task is to decide whether the THIRD word (X)
is the same as the FIRST WORD (A) or the SECOND WORD (B)
Press SPACE to continue


If the THIRD word is the same as the FIRST word
press the button on the LEFT
Press SPACE to continue


If the THIRD word is the same as the SECOND word
press the button on the RIGHT
Press SPACE to continue


If no answer is given in 3 seconds
the next word will automatically appear
ALWAYS give an answer
Press SPACE to continue


Some EXAMPLES have been provided for practice
ANSWER AS FAST AS POSSIBLE
READY?
Press SPACE to continue

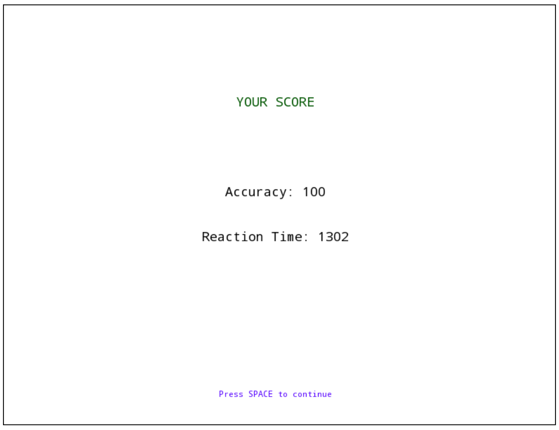

YOUR SCORE
Accuracy: 100
Reaction Time: 1302
Press SPACE to continue

Figure F.1. Instructions in the L2 vowel discrimination task.

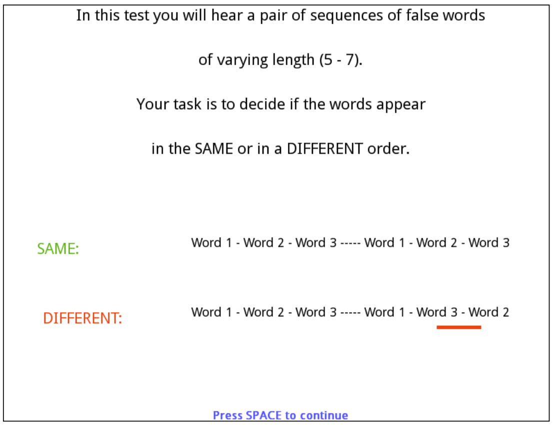
Appendix G. Phonological short-term memory task materials

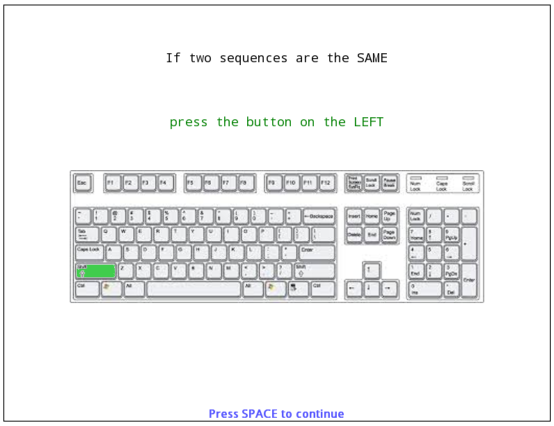
Table G.1

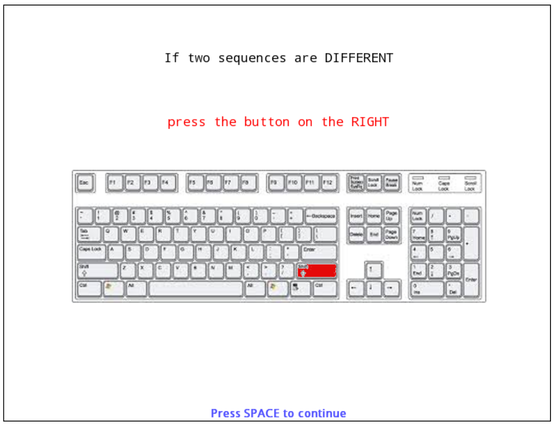
Stimuli in the phonological short-term memory task (from Cerviño-Povedano and Mora, 2011).

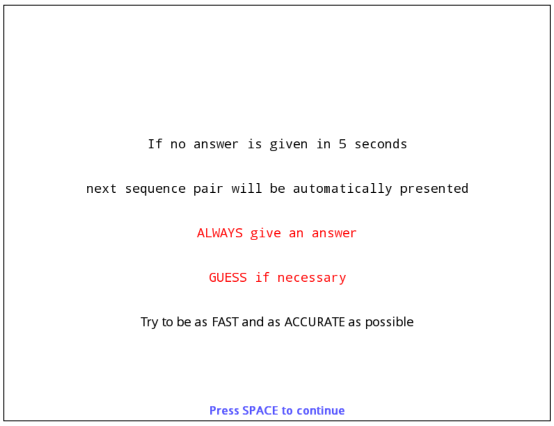
Practice block		1	2	3	4			
	1	ner	taʃ	bon	tiʎ			
	2	tɛm	**keʎ**	**fur**	ʒit			
	3	ʒis	kɛt	ʒul	maɲ			
	4	nut	**lim**	**lɔp**	neʎ			
5-item length sequences		1	2	3	4	5		
	1	ʎip	poʧ	kɛs	ʎun	ʃan		
	2	geʧ	ʎat	ʧɔl	ʒik	sum		
	3	soʃ	mir	nuʧ	ɲal	kɛr		
	4	nip	seʎ	ʒun	daʎ	pɔʧ		
	5	nal	**sɛm**	**bok**	mik	ʃul		
	6	ʒut	ɲal	**toʧ**	**teʎ**	bir		
	7	gɛn	**gur**	**diɲ**	meʎ	moʎ		
	8	keʎ	kut	**nɔl**	**lis**	ʎal		
6-item length sequences		1	2	3	4	5	6	
	9	ʎɛm	toʧ	bun	ʒes	ʎin	ɲɔl	
	10	gor	teʃ	bɔr	ɲin	paʎ	rɛn	
	11	bur	gɛʧ	ʎan	soʎ	fer	biɲ	
	12	ris	tas	sep	toʃ	ʃɛt	kuʃ	
	13	for	**ʎɛn**	**sir**	dul	teʧ	ʒɔl	
	14	pɛɲ	bor	**dup**	**dap**	meʃ	dil	
	15	ʎɔs	ler	rin	**tɛʃ**	**ʒan**	rup	
	16	raʎ	**gis**	**lek**	ʎup	soʧ	ʎɛs	
7-item length sequences		1	2	3	4	5	6	7
	17	doʃ	maɲ	pɔs	ɲɛm	ʒit	reʧ	kum
	18	fek	puʎ	roʎ	git	ɲat	gɛm	fɔʎ
	19	lat	rɔs	muk	tɛʧ	siɲ	fok	pes
	20	doɲ	tɔr	ken	taʃ	mip	kun	kɛm
	21	dɔs	lɛm	ʎik	**ges**	**kop**	nup	lan
	22	ger	tar	doʧ	sɛʧ	**kuɲ**	tɔs	biʎ
	23	fus	**met**	**fik**	lop	bɛm	faʎ	fɔl
	24	biʃ	lɛk	**fal**	**dɔp**	lep	kuʧ	top

Note: Items in bold are transposed items.


In this test you will hear a pair of sequences of false words
of varying length (5 - 7).
Your task is to decide if the words appear
in the SAME or in a DIFFERENT order.
SAME:
Word 1 - Word 2 - Word 3 ----- Word 1 - Word 2 - Word 3
DIFFERENT:
Word 1 - Word 2 - Word 3 ----- Word 1 - Word 3 - Word 2
Press SPACE to continue


If two sequences are the SAME
press the button on the LEFT
Press SPACE to continue


If two sequences are DIFFERENT
press the button on the RIGHT
Press SPACE to continue


If no answer is given in 5 seconds
next sequence pair will be automatically presented
ALWAYS give an answer
GUESS if necessary
Try to be as FAST and as ACCURATE as possible
Press SPACE to continue


The following items are for PRACTICE
4 non-word length sequences
READY?
Press SPACE to continue

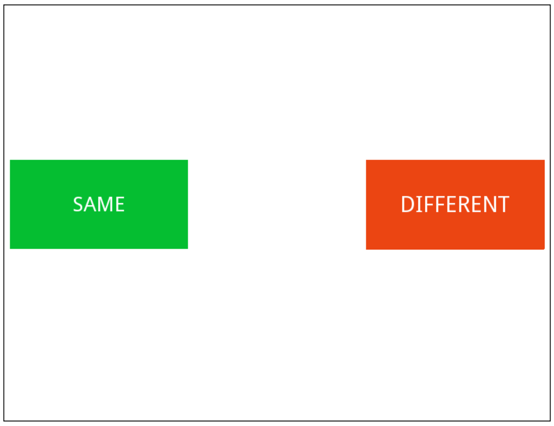

SAME
DIFFERENT

Figure G.1. Instructions in the phonological short-term memory task.

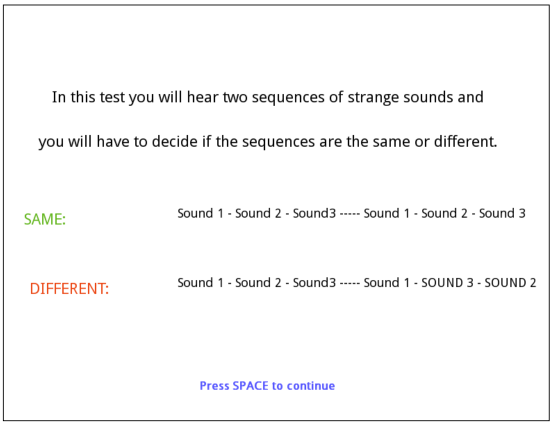
Appendix H. Acoustic memory short-term task materials

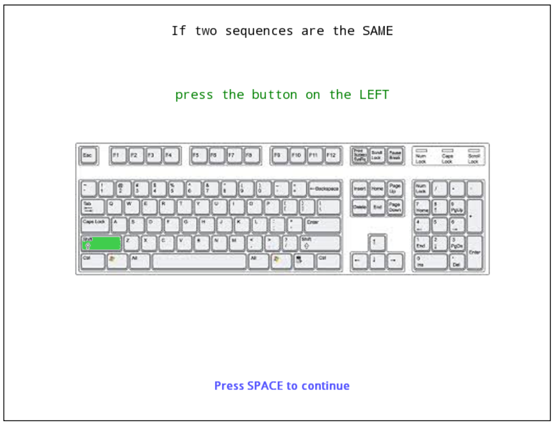
Table H.1

Stimuli in the acoustic short-term memory serial sound recognition task.

Practice block					
1	fi	ʎu			
2	**ʃe**	**do**	ɲi		
3	su	ʤi	po	li	
4	ma	re	**bu**	**si**	go
Block I: Three-item sequences					
1	ke	ʃa	ni		
2	ʎi	to	ba		
3	ze	ku	ri		
4	gi	ʎu	sa		
5	po	**le**	**ʃu**		
6	**da**	**mo**	si		
7	**ru**	**ʃi**	go		
8	so	**be**	**ʧa**		
Block II: Four-item sequences					
9	pu	li	fo	ra	
10	me	zu	ti	no	
11	bu	ta	ge	ʃo	
12	te	ʒu	ɲi	ga	
13	**ki**	**ʎa**	fu	ne	
14	ʧe	**ɲu**	**pi**	ʒo	
15	ka	**ʒi**	**lo**	de	
16	su	do	**fa**	**ɲe**	
Block III: Five-item sequences					
17	ʎe	mu	ʤi	ko	za
18	fe	ʎo	pa	ʧu	di
19	mi	zo	la	tu	ʤe
20	ɲa	pe	lu	zi	ʧo
21	**se**	**ʤu**	na	ri	bo
22	du	**ma**	**ʧi**	ro	ʒe
23	ʤa	re	**gu**	**fi**	ɲo
24	bi	ʤo	ʃe	**nu**	**ʒa**

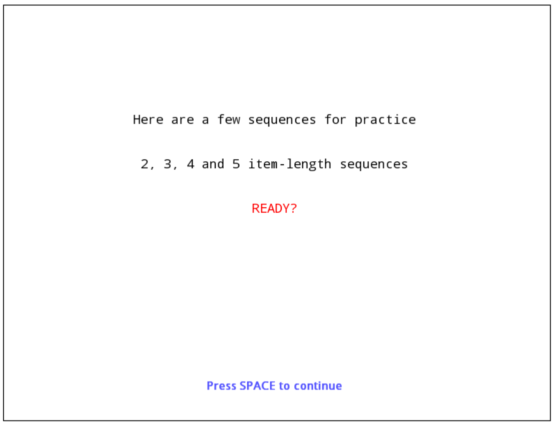
Note: Items in bold are transposed items.

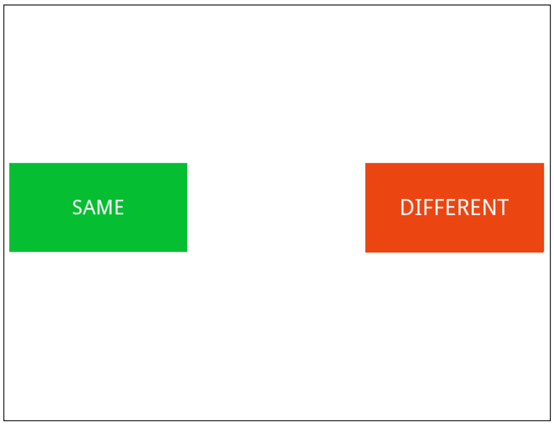

In this test you will hear two sequences of strange sounds and
you will have to decide if the sequences are the same or different.
SAME:
Sound 1 - Sound 2 - Sound3 ----- Sound 1 - Sound 2 - Sound 3
DIFFERENT:
Sound 1 - Sound 2 - Sound3 ----- Sound 1 - SOUND 3 - SOUND 2
Press SPACE to continue


If two sequences are the SAME
press the button on the LEFT
Press SPACE to continue

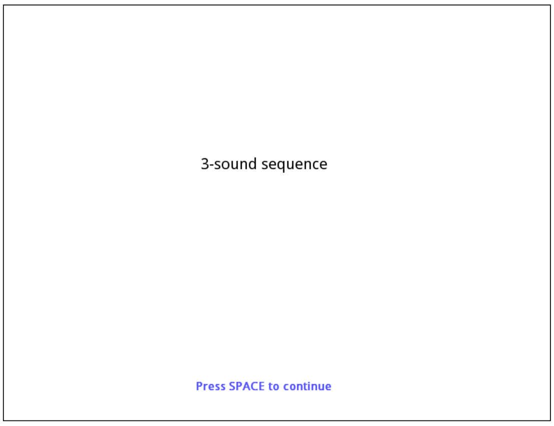

If two sequences are DIFFERENT
press the button on the RIGHT
Press SPACE to continue

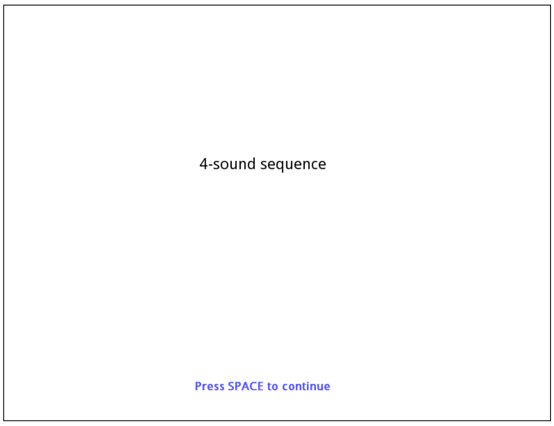

If no answer is given in 5 seconds
next sequence pair will be automatically presented
ALWAYS give an answer
Try to be as FAST and as ACCURATE as possible
Press SPACE to continue

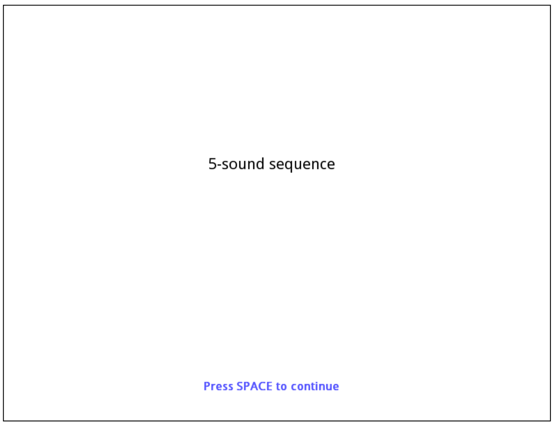

Here are a few sequences for practice
2, 3, 4 and 5 item-length sequences
READY?
Press SPACE to continue

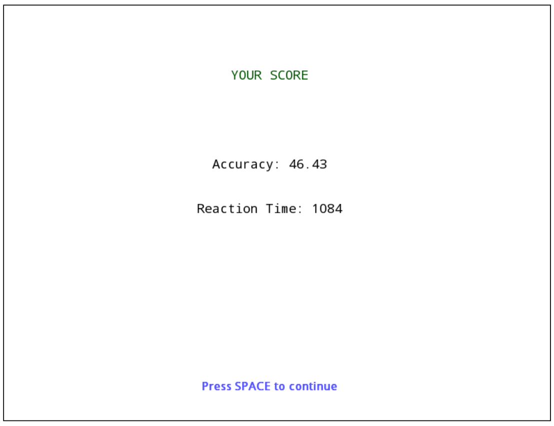

SAME
DIFFERENT


NOW THE TEST WILL BEGIN!
READY?
PRESS SPACE TO START


3-sound sequence
Press SPACE to continue


4-sound sequence
Press SPACE to continue


5-sound sequence
Press SPACE to continue


YOUR SCORE
Accuracy: 46.43
Reaction Time: 1084
Press SPACE to continue


THIS IS THE END OF THE TEST
THANK YOU FOR YOUR PARTICIPATION
PLEASE CALL THE RESEARCHER

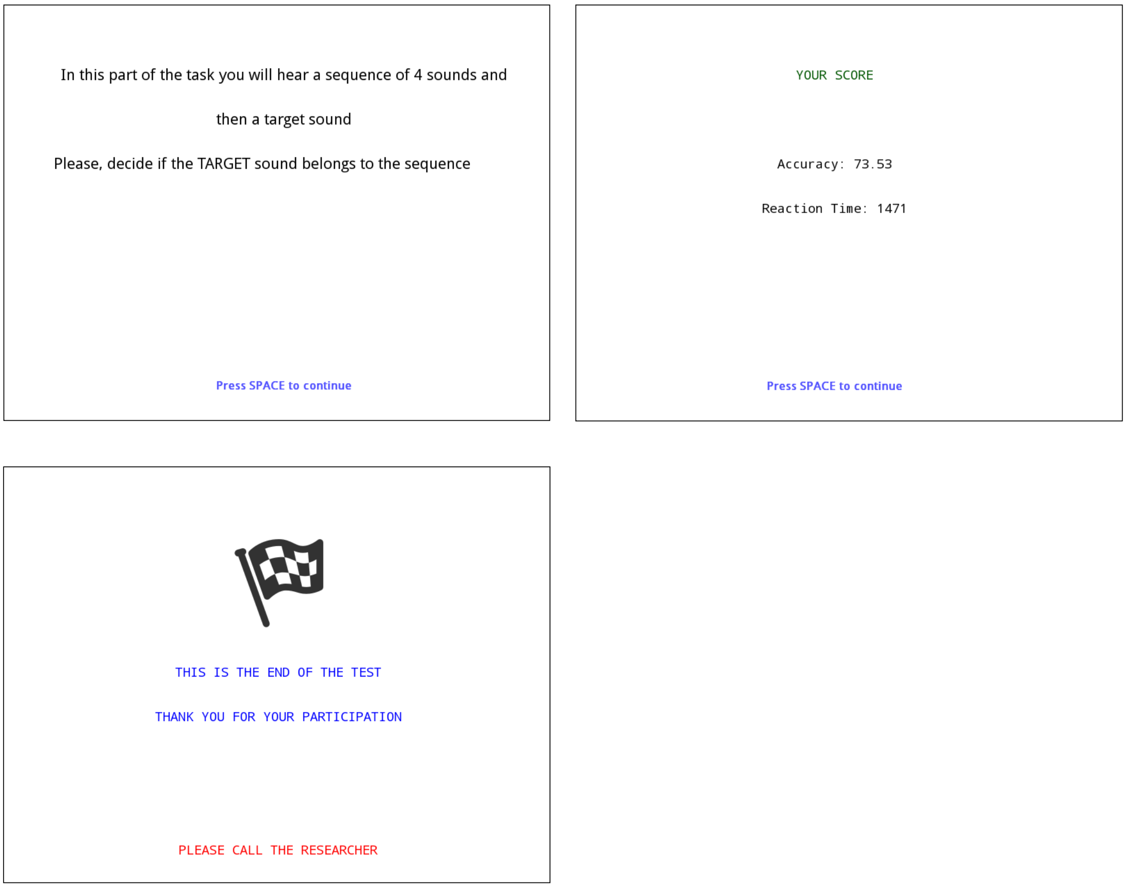


Figure H.1. Instructions in the acoustic short-term memory serial sound recognition task.

Table H.2

Stimuli in the acoustic short-term memory target sound recognition task.

Practice block						
1	**ʎu**	ɾe		**ʎu**		
2	ʃe	nu		ʒo		
Block I: Two-item sequences						
1	**zi**	ɲo		**zi**		
2	fu	**ge**		**ge**		
3	**ʒa**	ɾi		**ʒa**		
4	na	**tu**		**tu**		
5	me	ɾo		za		
6	lu	pi		ʧo		
7	pe	ʒa		ru		
8	ta	ʤo		ʎe		
Block II: Three-item sequences						
9	**la**	ke	ʤi		**la**	
10	ʎi	**to**	ba		**to**	
11	ze	ku	**mi**		**mi**	
12	ti	**ga**	ʃu		**ga**	
13	po	li	ru		ɲa	
14	da	mu	si		ʒe	
15	ɾu	ʃi	do		pa	
16	so	ʧa	be		ri	
Block III: Four-item sequences						
17	**fo**	le	pu	gi		**fo**
18	ɾa	**gu**	ʧi	mo		**gu**
19	bu	ne	ʃo	**ka**		**ka**
20	ɲi	ʒu	**te**	ma		**te**
21	ʒi	sa	ɲu	de		bo
22	ʧe	du	bi	zo		ra
23	ʎa	fi	go	se		ʤu
24	zu	fa	ko	ɲe		di

Note: Items in bold are the same items.

In this test you will hear a sequence of strange sounds and

after a short pause you will hear a TARGET sound which may

or may not belong to the sequence .

You should answer YES if the TARGET sound belongs to the sequence

YES: Sound 1 - Sound 2 - Sound3 Sound 2

You should answer NO if the TARGET sound doesn't belong to the sequence

NO: Sound 1 - Sound 2 - Sound3 Sound 4

Press SPACE to continue

If the TARGET sound is the SAME as one of

the sounds in the previously heard sequence

press the button on the LEFT

Press SPACE to continue

If the TARGET sound is the DIFFERENT from

the sounds in the previously heard sequence

press the button on the RIGHT

Press SPACE to continue

First, the following picture will appear on the screen

LISTEN TO THE SEQUENCE

Listen and try to remember all sounds in the sequence

Press SPACE to continue

Next, a picture of a loudspeaker will appear indicating

that the TARGET sound will be presented shortly

YES NO

And you will have to decide if the TARGET sound

belongs to the sequence or it doesn't

Press SPACE to continue

If no answer is given in 5 seconds

next sequence will be automatically presented

ALWAYS give an answer

Try to be as FAST and as ACCURATE as possible

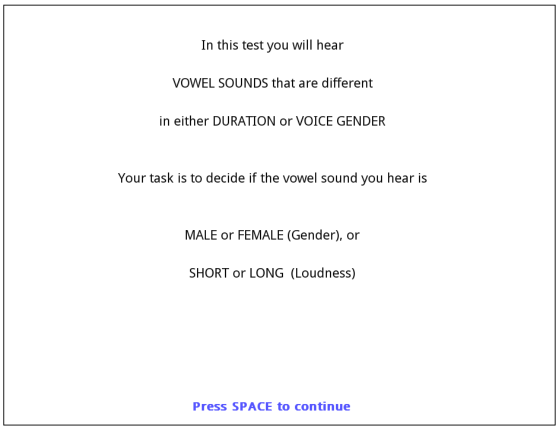
Press SPACE to continue

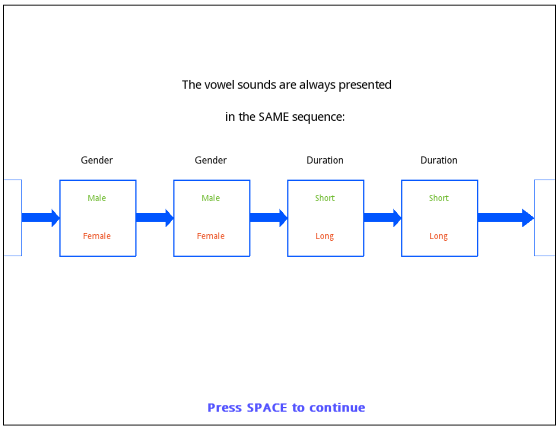
Figure H.2. Instructions in the acoustic short-term memory target sound task.

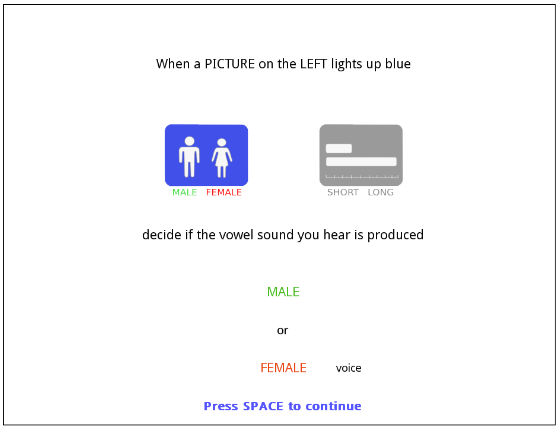
Praat script for spectral rotation

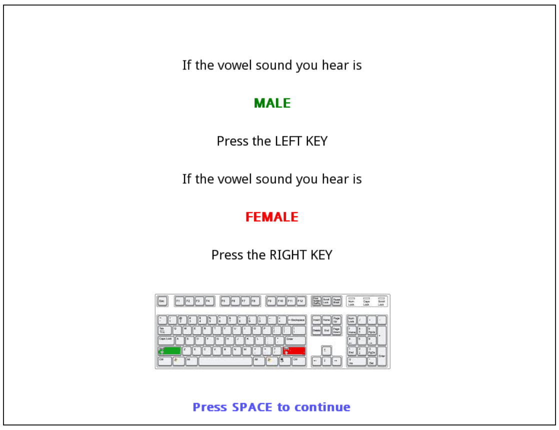
```
# ====================================================================
# rotation script for Praat
# low-pass filtered rotated version
# ====================================================================

form Rotate Sound
real Rotation_frequency 2000
word Name_of_new_sound rotated
endform

Copy... temp
d = Get total duration
sf = Get sampling frequency
Create Sound from formula... sine Mono 0 'd' 'sf' 1/2*sin(2*pi*'
    rotation_frequency'*x)
select Sound temp
Formula... self * Sound_sine[col]
select Sound temp
Filter (pass Hann band)...  0 3800 20
Rename... 'name_of_new_sound$'

select Sound temp
plus Sound sine
Remove
```

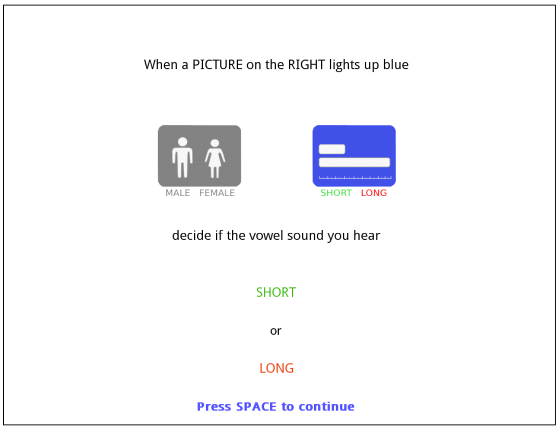
Appendix I. Attention-shifting task materials

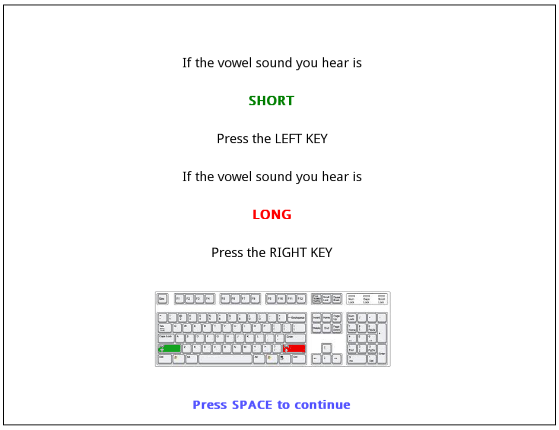












Reply as FAST and as ACCURATE as possible

GUESS if necessary

ALWAYS keep your HANDS on the KEYS!

You must PRESS A KEY within 3 seconds

Press SPACE to continue

Practice Block 1

Decide whether the vowel is

MALE (left)

or

FEMALE (right)

Press SPACE to start

This is the end of the practice block

Accuracy: 83.33

Reaction Time: 695

Press SPACE to continue

Practice Block 2

Decide whether the vowel is

SHORT (left)

or

LONG (right)

Press SPACE to start

This is the end of the practice block

Accuracy: 83.33

Reaction Time: 966

Press SPACE to continue

Practice Block 3

Decide whether the vowel is

MALE (left)

or

FEMALE (right)

SHORT (left)

or

LONG (right)

Press SPACE to start

Figure I.1. Instructions in the attention-shifting task.