Impact of the linguistic environment on speech perception: Comparing bilingual and monolingual populations

Abeba Roessler

TESI DOCOTRAL UPF / 2012

DIRECTORA DE LA TESI:

Dra. Núria Sebastián-Gallés (Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra)

UNIVERSITAT POMPEU FABRA

To my family

Acknowledgements

First of all, I would like to thank my supervisor Núria Sebastián-Gallés for guiding and supporting me throughout this dissertation. Through all ups and downs she has always been there helping me for example with complicated administrational situations and even making last corrections at 3 in the morning from a hotel room in America. She has not only taught me many important and valuable aspects about science but has also given me personal support in many ways.

Furthermore, there are many people who have contributed in their own ways to help me through the process of this dissertation. I want to thank Juan Manuel Toro for all the advice and help with running experiments and interpreting difficult results, Albert Costa and Clara Martín for helpful comments and advice, Ansgar Endress for giving me some insights into Hungarian vowel harmony and Luca Bonatti for always finding the important question to ask and having so much patience. I would also like to thank Jasmin, Fatima, Barbara and Marina for having become really good friends with whom I have shared many great moments throughout the last four years. Many thanks to Volker, Yu, Alba, Anna, Judith, Alexandra and Miguel for many relaxing and motivational conversations over lunch, to Nano for always listening and giving reassuring medical advice, and to Gisela and Nicoló for inspiring and sometimes quite philosophical discussions.

I want to especially thank Sara, Carolina, Daniela and Joana for sharing not only the office but also much of the stress and nervousness of the last few weeks. I hope I did not make you all go crazy!

Last but not least I want to thank Xavi, Silvia, Cris, Florencia and Jaume for all the support and help with some of the most important aspects of this dissertation.

Abstract

The present dissertation set out to investigate how the linguistic environment affects speech perception. Three sets of studies have explored effects of bilingualism on word recognition in adults and infants and the impact of first language linguistic knowledge on rule learning in adults. In the present work, we have found evidence in three auditory priming studies that bilingual adults, in contrast to monolinguals have developed mechanisms to effectively overcome interference from irrelevant information in the speech signal. Preliminary results on toddlers indicate no differences in the recognition of mispronounced words between bilinguals and monolinguals. Additionally, knowledge about rules in the first language was shown to have an impact on general rule learning abilities, while we did not detect an influence of bilingualism in this process. In summary, we have found evidence for an impact of the linguistic environment on the processing of indexical variability in word recognition as well as on rule learning. Bilinguals seem to have adapted to increased variability in their daily speech environment. In addition, rule extraction from unknown language input was unaffected by those adaptations but influenced by linguistic knowledge.

Resumen

El principal objetivo de este trabajo fue investigar cómo el entorno lingüístico afecta la percepción del habla. Con esta finalidad se exploraron los efectos del bilingüismo en el reconocimiento de palabras y el efecto de la lengua materna en el aprendizaje de reglas. Para ver los efectos del bilingüismo en el reconocimiento de palabras se realizaron tres estudios en adultos explorando el efecto de facilitación por repetición, que mostraron que las personas bilingües han desarrollado mecanismos que les permiten minimizar las interferencias que ejerce la información irrelevante en la señal del habla. Por otro lado, se realizó un estudio con niños pequeños cutos resultados sugieren que no hay diferencias en el reconocimiento de palabras mal pronunciadas entre niños pequeños bilingües y monolingües. Respecto al efecto del conocimiento lingüístico de la lengua materna en el aprendizaje de reglas, se mostró que tiene un impacto en las habilidades generales para el

aprendizaje de reglas aunque no se ha detectó una influencia del bilingüismo en dicho proceso. En resumen, se ha mostrado que el bilingüismo minimiza los efectos negativos de la variabilidad en el reconocimiento de palabras. Los bilingües parecen haberse adaptado a una mayor variabilidad en su entorno de habla cotidiana. Por el otro lado, se ha visto que la capacidad para extraer reglas de una lengua desconocida no está afectada por estas adaptaciones si no que está influenciada por los conocimientos lingüísticos en la lengua materna.

Preface

Early studies on bilingualism have mainly focused on the negative influences of having to cope with two languages, but recently it has been discovered that remarkable adaptations to this particular situation enable bilinguals to overcome the challenges that are inherent to the increased processing load. Apart from the obvious advantage of being able to communicate with a broader range of people, better performance of bilinguals on many cognitive tasks reflects effective adaptations of the language processing system. Clearly, there are differences between different types of bi-and multilingualism depending on factors such as age of acquisition, the specific languages and the type of environment of exposure. Early and highly proficient bilinguals supposedly have adapted their language system from early on and thus, should show the largest effects when compared to monolinguals.

The language environment in Barcelona offers a very interesting possibility to study bilingualism. First of all, both Catalan and Spanish are official languages and education in both languages is offered to children already at the kindergarten level. Furthermore, many families speak the two languages and thus, many children are exposed to both Spanish and Catalan in their family environment. Courses at school and University are held in both languages and every student has to pass an exam proving high proficiency in many aspects of both languages to enter University. A particular characteristic of these two Romance languages is that they share many features. An estimated 70% of words can be considered to have cognate status and also grammar and syntax share several key aspects. Differences between Catalan and Spanish are mainly found in the phonological systems. Spanish has fewer vowels than Catalan and some consonants exist only in Spanish, while others are specific to Catalan. Furthermore, vowel reduction, which means that vowels in unstressed positions are reduced, only exists in Catalan. In summary, growing up in Catalonia usually means that two languages are present in every-day life from very early on and that bilinguals have to manage the two languages on both the speech production and perception level. Much research on the cognitive consequences of bilingualism has explored effects on speech production and non-linguistic cognitive abilities. However, how cognitive adaptations to bilingualism impact speech perception, as well as the effects of the increased level of variability in the auditory signal, have received little attention.

The current dissertation set out to investigate how increased processing loads during speech perception, due to higher variability in the signal, influence basic auditory word recognition. Furthermore, we wanted to explore early adaptations to the bilingual environment by studying pre-verbal infants and their capacity to cope with mispronunciations. Finally, we were interested in more general effects of lifelong exposure to the rules of the native language on the processing of a new language and how bilingualism comes into play at this level.

Table of Contents

List of figures

List of tables

1. GENERAL INTRODUCTION

The complex human language system has been studied extensively for a long time. Still, there are many open questions and different areas that have been explored in little detail. Recently, bilingualism and its consequences on language processing and non-linguistic cognitive abilities have been the focus of many researchers. So far it has been assumed that most of the effects that have been observed for bilinguals are based on adaptations originating in the speech production system. However, using two languages on a daily basis implicates not only the production of two languages, but it also involves perceiving and understanding utterances from two languages. Hence, while there are many studies looking at the consequences of bilingualism on speech production and general cognitive abilities, comparisons of bilingual and monolingual populations on speech perception have received little attention. In the following introduction we want to overview some of the literature on cognitive consequences of bilingualism and describe the situation bilingual listeners are confronted with in their daily life. We will argue that bilinguals are exposed to increased variability in their listening environment and describe how this may affect basic speech perception processes. Furthermore, we will discuss the known effects that have been observed in bilingual speech perception in adult listeners and infants. We will then go on to consider how the native language background may influence learning of a new language and how bilingualism interacts at this stage with knowledge from the first language.

1.1 Bilingualism and speech perception in adults

1.1.1 Cognitive consequences of bilingualism

Differences between bilinguals and monolinguals have been detected from very early on during language development. For instance, only bilingual infants were able to extract more than one rule from an artificial speech stream (Kovács & Mehler, 2009b) and they exceeded performance of monolingual peers at theory of mind (TOM) tasks (Kovács, Téglás, & Endress, 2010). Throughout childhood (Bialystok, 1999; Bialystok & Martin, 2004) and as adults evidence showed that bilinguals performed better on switching tasks (Prior & Macwhinney, 2010), Simon tasks (Bialystok, Klein, Craik, & Viswanathan, 2004) and Stroop tasks (Bialystok, Craik, & Luk, 2008a), which are tasks that required executive control. Parallel activation of both languages during speech production, which has been demonstrated in several studies, has been suggested to be at the base of these findings, because selection of the appropriate language is necessary for producing coherent utterances (e.g. Colomé, 2001; Costa, La Heij, & Navarrete, 2006; Hernandez, Bates, & Avila, 1996; Martin, Dering, Thomas, & Thierry, 2009). There is still an intense debate about the exact nature of the mechanism that enables bilinguals to deal with parallel activations during speech production (see Costa, 2005 for a discussion). Converging evidence however, suggests inhibition (interference suppression) as the main mechanism. Irrelevant information is inhibited to varying degrees depending on the task context, which leads to an advantage for bilinguals in ignoring irrelevant information as well in other tasks (Bialystok, 2009 for a review; Philipp, Gade, & Koch, 2007; Philipp & Koch, 2009; Treccani, Argyri, Sorace, & Della Sala, 2009). Another mechanism that has been considered to be more effective in bilinguals is related to the fact that bilinguals need to constantly monitor and manage two languages A bilingual needs to keep track of the language in use and adjust quickly to the given circumstances, for example if a new person enters a conversation. Hence, this training may translate into better capacities in keeping track of task sets and rapid correct selection of the target. Evidence supporting this view has been provided by several studies looking at attentional blink (Philipp & Koch, 2009), the flanker task (Costa, Hernandez, Costa-Faidella, & Sebastián-Gallés, 2009; Costa, Hernández, & Sebastián-Gallés, 2008) and other paradigms that use congruency (Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés, 2010). For instance, when performing a task, which includes congruent and incongruent trials, bilinguals are overall faster and better at responding to both types of trial, although no conflict exists in congruent trials. This has been interpreted as reflecting capacities to monitor task requirements, which may stem from practise in monitoring the language in use.

Apart from the studies on general cognitive abilities, most research on bilingualism has focused on reading (on the comprehension side) and on speech production processes, whereas only very few studies have looked at how bilinguals deal with variable speech input on the auditory level and how this may influence general speech perception mechanisms.

1.1.2 Bilingual speech perception

Although most of the effects of bilingualism that have been reviewed so far are attributed to the word selection problem during speech production, where only one word candidate and language have to be selected to produce coherent utterances, the increased level of variability during auditory speech perception in a bilingual environment could have some important implications as well. In an environment where two languages are present, it is often the case that listeners are frequently encountering non-native speakers who produce accented or mispronounced utterances. For bilingual populations, including the one used throughout the dissertation, each of the two languages has its own phonology and some sounds only exist in one language. It is therefore highly likely that those speakers producing words in their non-native language have difficulties in pronouncing them correctly if they involve a non-native phoneme. Thus, to effectively understand utterances from a large percentage of possible interlocutors, a bilingual has to take into account that many variations in pronunciation are possible. In fact it has been shown that bilinguals retain accented versions of words in their lexicon (Sebastián-Gallés, Vera-Constán, Larsson, Costa, & Deco, 2009) and that when confronted with mispronounced utterances bilinguals are more flexible at accepting or recognising those mispronunciations (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009; Sebastián-Gallés, Echeverria, & Bosch, 2005; Sebastián-Gallés, Rodríguez-Fornells, De Diego-Balaguer, & Díaz, 2006; Sebastián-Gallés, et al., 2009). These results point toward an increased number of activated entries during bilingual auditory speech perception at least when involving accented and mispronounced utterances.

Furthermore, several studies on lexical activation have found that bilinguals access their two lexicons in parallel not only during speech production, as mentioned earlier, but also during speech perception. In a series of studies by Marian and colleagues (Marian, Blumenfeld, & Boukrina, 2008; Marian, Spivey, & Hirsch, 2003; Spivey & Marian, 1999) clear effects of parallel activation of both lexicons were observed when bilinguals were listening to auditory stimuli in one of their languages. Using eye-tracker measurements, within-language as well as between-language competitor fixations were observed, showing that at least in the early stages of word recognition both languages were active (Marian, et al., 2008; Marian & Spivey, 2003).

In addition to the increased processing load described so far (in terms of activation of more lexical candidates), it has been reported that "spurious" activation occurs during bilingual speech perception. Pallier, Colomé & Sebastián-Gallés (2001), presented Spanish-Catalan bilinguals with minimal pairs using Catalan-specific contrasts that were generally difficult for Spanish natives to perceive. Repetition priming effects were observed in Spanish native bilinguals, thereby showing that due to insensitivity to these contrasts, Spanish native bilinguals activated those items as homophones.

Using a visual world paradigm, voice onset time (VOT) of Spanish words was manipulated to correspond either to Spanish or English pronunciation. Fixations to target pictures as well as to pictures, whose English translations were phonologically similar to the Spanish targets were explored. It was found that only when VOTs corresponded to English, would between language competitors receive more fixations. Results from this study suggest that when bilinguals are able to perceive fine grained acoustical detail, it can be used to restrict activations to the appropriate language (Ju & Luce, 2004).

In general, those studies suggest that, when listening to speech, bilinguals have to deal with more widespread activations, which are due to the non-selective lexical access, the higher number of accented and mispronounced utterances and spurious activation based on insensitivity to certain contrasts. It has to be noted as well that L2 phonology is usually not as well represented as L1 phonology and thus, lexical access is generally noisier. Therefore, some form of mechanism for sorting through those activations and selecting the appropriate candidate or some form of interference suppression is also likely to be necessary during auditory speech perception.

1.1.3 Spoken word recognition and input variability

Auditory processing is incremental in nature and involves high levels of variability in the surface form of the signal, due to different acoustical characteristics of for example the particular speaker. Thus, the question of how the system deals with the great variability in the input in auditory processing has not been completely resolved to date. Understanding messages from different speakers with different voice qualities, accents and dialects, is no real problem for the brain, while so far it has not been possible to recreate this efficiency computationally. Traditionally, it has been assumed that abstract representations are stored in the mental lexicon and input has to be normalised, stripping off all indexical, irrelevant information as the source of noise (Luce & Pisoni, 1998; Norris, 1994). A large body of evidence supports abstraction, but recent studies have shown that this cannot be the whole story. In a variety of studies it has been found that fine grained detail can be preserved in long-term memory, with effects for example on priming, which indicates automatic processing of details (Bradlow, Nygaard, & Pisoni, 1999; Church & Schacter, 1994; Goldinger, 1996; Nygaard, Sommers, & Pisoni, 1995; Palmeri, Goldinger, & Pisoni, 1993; Schacter & Church, 1992).(Sommers et al., 1994; Mullenix et al., 2002). Most of these studies showed that repetition priming was disrupted or completely abolished when indexical features (such as the speaker or the accent) changed between the first presentation and the repetition. This reduction in priming has been observed in several studies, but results are inconsistent and seem to depend strongly on the specific task and testing conditions (González & McLennan, 2007; Luce & Lyons, 1998; Orfanidou, Davis, Ford, & Marslen-Wilson, 2011; Orfanidou, Marslen-Wilson, & Davis, 2006). Due to the fact that replicating those results has proven difficult, recent accounts suggest a more complex model of word recognition, where abstract and detailed episodic representations interact to create the behavioural output that has been observed. Several hypotheses have been proposed, that could explain the results. Some authors have suggested that the time-course is of importance, in the sense that slow responses are influenced by slow acting episodic traces, while fast answers are mainly based on abstract representations (McLennan & Luce, 2005;

Orfanidou, et al., 2011). Alternatively, it has been proposed that indexical (episodic) information is stored outside the lexicon and does not necessarily influence word recognition (Bowers & Kouider, 2003).

In general, recognition memory paradigms are supposed to give rise to episodic effects during word recognition, because they involve slow and rather effortful processes (McLennan & Luce, 2005). In such a paradigm participants have to respond if a word they hear is “new” (presented for the first time) or “old” (repeated), irrespective of the voice producing the word. In contrast, in a lexical decision task responses are fast and can be based on abstract knowledge, so responses should be less affected by indexical information.

In the present work, we used both recognition memory and lexical decision tasks to compare bilinguals and monolinguals on their processing of irrelevant speaker information

1.1.4 The present study

As reviewed above bilinguals have to deal with two languages in addition to all challenges a monolingual listener is faced with. It is assumed that bilinguals have recruited the executive function system to a larger degree to accomplish the additional processes involved in managing two languages. Consequently, advantages on tasks requiring executive function resources have been observed. However, it is still unclear what kind of consequence may be provoked by the increased levels of variability in the daily speech input of bilinguals. The fact that bilinguals have to accommodate accents and mispronunciations on a daily basis as well as the increased processing load caused by the multiple, language non-specific activations during speech perception, may have as a consequence the development of a mechanism that helps to rapidly deal with irrelevant information during speech perception. Thus, using speaker variability as irrelevant dimension we wanted to test if bilinguals could deal more efficiently with indexical variability.

To test the hypothesis that, due to adaptations to the increased input variability, bilinguals are less affected by irrelevant indexical information, we ran several auditory priming studies varying the task and indexical information at the speaker level and compared bilingual and monolingual populations. Specifically, we used two recognition memory tasks with varying degrees of variability (2 or 8 speakers) to tap into processing of irrelevant speaker information and possible differences between bilinguals and monolinguals in overcoming interference. Furthermore, we conducted a lexical decision task to explore the differential influence on bilinguals and monolinguals of indexical variability in the environment, when processing words and non-words. With this second paradigm we may be able to entangle specific effects of speaker variability in word recognition from general effects of variability in the environment (Chapter 1).

1.2 Bilingualism and speech perception during early language development

1.2.1 Effects of bilingualism on young language learners

As mentioned earlier, most results on cognitive advantages in bilinguals have been attributed to adaptations necessary for speech production. However, the fact that pre-verbal bilingual infants at 7 months of age displayed enhanced cognitive abilities, compared to monolinguals, suggests that some adaptations may not be based on speech production (Kovács & Mehler, 2009a). Several potential sources of early effects are possible. Apart from the fact that there are two sets of rules, phonology and vocabulary to be recognised and learned, accommodating higher levels of variability during speech perception may also exert an early effect. Studying young bilingual language learners could tap into effects stemming from speech perception rather than speech production.

Considering the particular example of Spanish and Catalan, many words are cognates, which differ mainly in regard to vowels. Meaning that consonants stay relatively constant and vowels are a more variable unit between the two languages. It is therefore possible that for Spanish-Catalan bilinguals one adaptation to the language input may be paying more attention to consonants and less to vowels, when compared to monolinguals.

This idea is supported by evidence from infant studies carried out on Spanish-Catalan bilinguals (Bosch & Sebastián-Gallés, 2003; Sebastián-Gallés & Bosch, 2009). As part of those studies, vowel processing in bilinguals during language development was explored. It is well known that after birth infants have a (quite) universal ability to perceive most phonetic contrasts (Polka &

Werker, 1994). In the first year of life however, the input language shapes the perceptual space and by the end of the first 12 months infants have formed categories for their native phonemes, while losing the ability to perceive some foreign contrasts (Kuhl, 2004; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994; Werker & Tees, 2005). This perceptual tuning process has been extensively studied in monolingual babies from different language backgrounds, but what happens in bilingual babies was an open question until fairly recently. In a first experiment the Catalan contrast between /e/ and /ɛ/ was tested in monolingual Spanish infants, monolingual Catalan infants and finally in Spanish-Catalan bilingual babies. This was done at 4 and 8 months of age. As expected at 4 months all tested groups showed discrimination of the critical contrast, as native tuning had not yet shaped the system. At 8 months this process should be well underway and accordingly Spanish monolingual babies were not succeeding in the task, while Catalan monolingual babies, for whom the contrast is native, did discriminate. Surprisingly however, the bilingual babies did not show discrimination, although they had been exposed to this contrast. In a follow-up at 12 months it was found that bilingual infants did discriminate those contrasts after all (Bosch & Sebastián-Gallés, 2003). This could be interpreted as evidence for a delayed development of phonetic categories or another possibility is that categories for /e/ and /ɛ/ have merged into one /e/ category as this is the most frequent version to be encountered between both languages. To further elucidate these surprising results, other contrasts have also been tested. When comparing the /o/-/u/ contrast that is common to both languages, the same results were obtained, indicating that bilingual babies did not discriminate. Only the common and distant contrast between /e/ and /u/ did finally result in discrimination behaviour of 8-month-old bilingual infants (Sebastián-Gallés & Bosch, 2009). In contrast, similar studies carried out in Canada using English-French bilinguals have produced ambiguous results. While some studies also indicate delayed tuning for bilinguals (Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006), other studies find no differences in processing between monolingual and bilingual infants (Burns, Yoshida, Hill, & Werker, 2007; Sundara, Polka, & Molnar, 2008). Thus, it seems that conflicting evidence makes it hard to interpret those findings as there are usually differences in the task (requiring attention or not) and materials

used (vowels vs. labial and dental consonants), which could contribute to the results.
All studies mentioned so far have used a task in which the infants are habituated to one type of stimulus (e.g. /dɛdi/) and then tested on another stimulus using a specific contrast (e.g. /dedi/). This type of paradigm is based on the assumption that infants recapture attention if they detect a change between habituation and test (Jusczyk & Aslin, 1995). Importantly, using an anticipatory eye-movement task, which is not based on recapturing attention, Albareda-Castellot, Pons, & Sebastián-Gallés (2011), tested the same /e/-/ɛ/ contrast and found that bilingual infants at 8 month of age were able to discriminate. This study has shown that the failure to observe discrimination in bilinguals in the previous studies was due to an attentional factor, meaning that bilingual infants were less surprised than monolinguals by a change in vowels between /e/ and /ɛ/ or /o/ and /u/. Taken together these studies suggest that Spanish-Catalan bilingual infants are not delayed in their native language tuning, but rather pay less attention or are less surprised by changes in certain vowels. Given the discrepancies between studies on Spanish-Catalan bilinguals and those conducted with English-French bilinguals, it can be argued that this is due to the specific attributes of Spanish and Catalan. In contrast to English and French, they share a great number of cognate words, as mentioned above. When taking 150 words from the McArthur-Bates inventory for infants, an estimated two thirds of the words can be classified as cognates (unpublished data Pons & Sebastián-Gallés). It also becomes quickly apparent that in most cases, cognates share the consonants, while vowels vary (*e.g.* pelóta - pilɔ́tə). And as children at this age are doing extremely well at tracking regularities and finding structures in their input it might be a useful strategy for Spanish-Catalan bilingual babies to pay more attention to consonants than to vowels, as these are usually the common structure between the two languages. Also, this could explain why in Bosch & Sebastián-Gallés, (2003) and Sebastián-Gallés & Bosch, (2009) infants were not surprised by the change in vowels, as this happens frequently in their environment. Taken together those studies point toward more flexibility of bilinguals in accepting small phonetic changes during speech perception. In addition, in the same way as adult bilingual listeners, infants are exposed to increased levels of variability in the pronunciation of words. As explained before many speakers are producing mispronounced

utterances when speaking in their non-native language and especially similar words may be subject to errors in the correct pronunciation.

1.2.2 Accent processing by young language learners

To discover and learn the native language(s), infants need to acquire many complex rules and principles. Apart from many others, phonological distinctiveness as well as phonological constancy are crucial features for speech perception. Minimal word pairs differing only in one phoneme have to be identified as two separate items, while words have to be recognized as the same item across different pronunciations. Several studies have explored infants´ recognition of accented or mispronounced words in a variety of tasks.

Best, Tyler, Gooding, Orlando, & Quann, (2009) used a familiarity preference task presenting familiar and unfamiliar words to infants of 15 and 19 months of age. Infants showed a preference to listen to familiar words in their native dialect (Connecticut English). When words were presented in a foreign dialect (Jamaican Mesolect) this familiarity preference disappeared for 15-month-old infants, suggesting that words were not recognized when spoken in the foreign accent. 19-month-olds however showed the familiarity preference also for words spoken in the Jamaican Mesolect, providing evidence for the emergence of phonological constancy at around 19 months of age. Using a different task based on familiarisation, Schmale & Seidl, (2009) tested 9 and 13 month old infants. Familiarisation was either spoken by one of two native English speakers or one of two Spanish accented speakers. When tested, 13 month olds could recognise native and accented test passages for both native and accented familiarisation, while at 9 months infants only recognised accented items when previously familiarised with the same Spanish accented speaker. This shows that infants as young as 13 months of age, and under some circumstances even 9 month-olds, can deal with accented speech in a familiarisation task

In a different approach, White & Aslin (2011) used a systematic shift in a specific vowel instead of natural accents. Infants at 18 - 20 months of age were exposed to either standard or shifted pronunciations, while viewing the corresponding objects. In a test phase, familiar objects were presented with new objects and infants’

word identification was tested using both standard and shifted pronunciations. Exposure to the shifted pronunciation was necessary for infants to be able to identify objects successfully, when labelled with the shifted pronunciation at test.
Considering results from the reviewed studies, it seems that infants as early as 9 months of age can show recognition of accented speech in a familiarisation task (Schmale & Seidl, 2009). However, using a word recognition task, Best, et al., (2009) found that only 19 but not 15-month-old infants showed a familiarity preference for known words in an unknown dialect. It has been concluded that around 19 months of age the concept of phonological constancy emerges, although vocabulary size has been found to be a better predictor of the recognition performance on accented speech than age. Importantly, the role of exposure to mispronunciations was demonstrated by White & Aslin, (2011). For word identification (relate the label to a picture) infants needed exposure to the "dialect" to perform the task successfully. Therefore, those studies give indirect support to the idea that the constant exposure to many variable pronunciations in the bilingual environment may have as a consequence that bilingual infants are more flexible at accepting mispronunciations.

1.2.3 The present study

In the present research, we set out to investigate possible adaptations of bilinguals to deal with increased processing loads. The exposure to an environment with high levels of mispronunciations may have lead to more flexibility at the time of word identification. In addition, due to the shared structure of cognate items, common in Catalan-Spanish bilingual environments, it is possible that bilingual infants exposed to these two languages pay less attention to variable vowels and thus, are less influenced by a shift in vowels during word identification. To test this hypothesis we used the same design as White & Aslin, (2009) and tested bilinguals and monolinguals on their ability to identify objects based on standard and shifted pronunciations. We only included the condition without previous exposure to the shifted versions to specifically test if living in the bilingual environment has provided enough exposure to shifted pronunciations to perform successfully on this type of word identification task (Chapter 2).

One way that infants could learn to pay more or less attention to certain features of their input is by tracking statistical regularities. Infants are particularly able to extract statistical information from their environment in many different domains.

1.3 Influence of language background on extraction of statistical regularities

So far, we have given an overview of how speech perception can be influenced and guided by exposure to the native language(s). But what exactly happens during exposure to language? Recently it has been suggested that intrinsic abilities to use fast and automatic mechanisms to extract relevant information from the environment play a crucial role in acquiring language.

It has been shown that computing statistical regularities happens across different domains, such as during visual or auditory processing (e.g. Endress, 2010; Fiser & Aslin, 2001, 2002; Jonaitis & Saffran, 2009). Importantly, statistical probabilities and regularities are computed during language processing to extract information about words and rules (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Peña, Bonatti, Nespor, & Mehler, 2002). In particular, an early study on 8 month old infants provided evidence that word boundaries could be extracted from fluent speech based only on statistical cues (Saffran, Aslin, & Newport, 1996). In a natural language setting, speech is perceived as a continuous stream, rather than separate words, because pauses in speech production do not correlate with word boundaries. For infants this means when listening to speech, statistical dependencies in the signal help to segment and analyse the input for regularities and to eventually extract words and rules of the input language.

But not only infants are faced with unknown speech input and evidence from studies on adults shows that the same mechanisms are involved in adult processing of foreign speech input. Thus, it seems that language learning is both based on the exposure to the specificities of the input language, as well as on the general intrinsic abilities to extract regularities and patterns from this input.

1.3.1 Statistical learning of words and rules in artificial languages

Many studies have used statistical learning paradigms to explore what type of regularities can be extracted from continuous speech. For instance, dependencies between consonants and syllables, besides other units, can be used to compute statistical regularities (Bonatti, Peña, Nespor, & Mehler, 2005; Mehler, Peña, Nespor, & Bonatti, 2006; Saffran, Newport, & Aslin, 1996). By creating an artificial language that was controlled for transitional probabilities between syllables it has been shown that also adult listeners are able to extract word boundaries based only on this kind of probabilities (Saffran, Newport, et al., 1996).

Many experimental approaches to statistical learning are based on artificial languages and can explore if participants are able to extract word-boundaries from a continuous speech stream, as well as learn a rule that is instantiated over vowels (Peña, et al., 2002).

Thus, to extract words, transitional probabilities for example between consonants can be computed to establish which syllables go together more often than others (Saffran, Newport, et al., 1996). In this type of paradigm, fixed consonant frames are combined with vowels that follow for example an ABA rule, resulting in trisyllabic (CVCVCV) non-words. By concatenating those non-words, a stream is constructed in which the probability between consonants within a word is higher than between words. A few minutes of exposure to such a stream is sufficient to compute these statistics and successfully extract word-boundaries and the vowel rule (Nespor, Peña, & Mehler, 2003; Toro, Nespor, Mehler, & Bonatti, 2008).

Different roles have been attributed to consonants and vowels in this process. By reversing the role of consonants (instantiate rule) and vowels (fixed frames), it has been shown that the task cannot be solved. This has been interpreted as evidence that consonants are the important unit across most languages that is used to extract information about word boundaries, while vowels give more information about syntax and rules (Nespor, et al., 2003; Peña, et al., 2002; Toro, et al., 2008).

1.3.2 Constraints of linguistic knowledge on statistical learning

However, as discussed before, characteristics of a known language can influence how speech is perceived, but so far, only few studies have looked at the influence of linguistic knowledge on the extraction of information from an artificial language.

Some studies on infants (Johnson & Seidl, 2009; Thiessen & Saffran, 2003) and adults (Fernandes, Ventura, & Kolinsky, 2007; Shukla, Nespor, & Mehler, 2007) provide evidence that some linguistic features such as prosody, stress and co-articulation can interact with statistical learning, thereby influencing word segmentation from continuous speech.

Furthermore, several studies have explored the impact of specific knowledge in the native language on statistical learning. Finn & Hudson Kam (2008) tested English native adults on their capacity to segment a continuous speech stream, using consonant clusters at word onset that were illegal in English. They found that this manipulation disrupted word segmentation and concluded that, indeed, restrictions in the first language impact the extraction of words from a new language.

Toro, Sebastián-Gallés, & Mattys, (2009) explored the influence of stress-placement on the word-initial, medial or final syllable in native listeners of French, Spanish and English. Due to the fact that stress location in those three languages differs, these groups may be affected by this manipulation as a function of stress placement in the native language. In French stress is placed on the final syllable and it is not considered a contrastive feature. In Spanish stress is usually on the penultimate syllable, while for English stress is mostly word initial. Importantly, in both English and Spanish minimal word pairs may only differ in stress placement and thus stress is contrastive in these languages. Surprisingly, both Spanish and English speakers were affected negatively, when stress was placed on the medial syllable, which may be due to the fact that attention was drawn away from word boundaries. French speakers were not affected by any manipulation, suggesting that if stress is not contrastive listeners may be deaf to those manipulations. Those results provide evidence for a general attentional effect of stress placement on word segmentation, as well as a language specific effect of the relevance of certain cues in the native language.

More recently, Toro, Pons, Bion, & Sebastián-Gallés, (2011) analyzed the influence of vowel reduction (a phonological rule constraining the subset of possible vowels appearing in stressed and unstressed positions. In chapter 3 a more exhaustive description is provided). These authors created a continuous speech stream including violations of Catalan word forming rules and showed that Catalan native listeners were impaired at selecting the appropriate word candidates, when those included violations of Catalan vowel reduction. However, if the possible candidates were tested against words that had not appeared in the stream or when they were displayed visually Catalan natives performed the task above chance. Those results suggest that statistical regularities were extracted but the selection of the appropriate candidate was significantly affected by restrictions in the native language. The three studies reviewed above provide both evidence for an impact of specific native language knowledge and universal linguistic constraints.

1.3.3 The present study

All those studies have demonstrated effects of previous knowledge on word segmentation processes. However, considering that the present dissertation has so far dealt with effects of bilingualism on speech perception processes, it could be interesting to explore on a more general level how exposure to a certain set of rules can affect processing of rules in a new language. To this end, we wanted to first test if rule learning can be similarly affected as word segmentation by first language knowledge. Therefore, we selected three groups of participants whose native language (Spanish, Catalan and Hungarian) differed in terms of the rules that applied to vowels at the word level and tested them on their capacity to extract a simple rule from an artificial speech stream. In a second step, we wanted to contrast possible effects of the native language with potential influences stemming from bilingualism. Thus, we tested an additional bilingual group, whose native language was Spanish, while having a good command of Catalan as well. Comparing this group to both Catalan native bilinguals (with a good command of Spanish) and Spanish monolinguals could reveal if potential effects are rather due to the rules of the first language or the fact that two sets of rules have been learned.

In the present work we wanted to extend findings on statistical learning to rule learning as so far all studies have looked at the influence of prior knowledge on word segmentation. In addition, we wanted to explore how bilingualism and first language knowledge may interact when faced with a new language.

1.4 General objectives

The present dissertation sets out to explore in several series of studies the effects of the linguistic environment on auditory speech perception. Testing bilingual and monolingual adults on their capacity to overcome indexical variability in the auditory input, due to increased variability in the daily environment, could provide new information about the effects of bilingualism, as well as shed more light on general speech perception mechanisms. Furthermore, adaptations by bilinguals to deal with the increased processing loads might start early in life and already affect word recognition in young infants. Thus, testing the ability of 19-month-old bilinguals to identify mispronounced depicted words, could reveal at which stage during language development bilingualism exerts its influence and how bilinguals adapt their speech perception system. Finally, we want to test several groups of adult listeners, whose native language differs in the way vowels are used. We want to compare their capacity to extract a new rule from ongoing speech, thereby exploring the more specific influences the linguistic environment imposes on subsequent speech processing.

The following chapter with the title *BILINGUALISM AND INDEXICAL VARIABILITY IN SPEECH PERCEPTION* and the authors Abeba Roessler, Beatriz Gil-Gómez de Liaño and Núria Sebastián-Gallés has been submitted for publication in *Cognition.*

2. BILINGUALISM AND INDEXICAL VARIABILITY IN SPEECH PERCEPTION

2.1 Introduction

Speaking and understanding more than one language has some obvious advantages in a world that becomes increasingly more global, but still most studies on speech perception and language processing have been conducted on monolingual populations. However, learning two (or more) languages poses very different challenges to the speech perception system and recently the study of bilingualism has received renewed attention. Differences between bilinguals and monolinguals on tasks involving executive function (Bialystok, Craik, & Luk, 2008b; Bialystok, et al., 2004; Costa, et al., 2009; Costa, et al., 2008) and speech production (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Roberts, Garcia, Desrochers, & Hernandez, 2002) have been found. Regarding speech production, many studies have observed that bilinguals produce fewer words and make more errors on verbal fluency tasks than monolinguals, which has been attributed to the increased vocabulary size. On the other hand, tasks tapping into executive function abilities (e.g. Simon and Stroop tasks) and switching paradigms have found better performance of bilinguals. It is generally assumed that practise in managing two languages, especially at the speech production level, lead to better general executive function abilities (for reviews see Bialystok, 2009; Bialystok, 2010). So far however, few studies have looked at the consequences of bilingualism on the processing of variable speech input (e.g. mispronunciations and accents). Consider a scientist going to a conference with many different non-native speakers presenting their work in English. It takes some time to get used to understanding this myriad of different pronunciations, but after years of listening to talks by non-native speakers listeners get really good and much more flexible at understanding almost any accent, while a person who never listens to foreign accents has difficulties understanding. Bilinguals however, are confronted with accented and mispronounced utterances on a daily basis and have gathered experience with variable speech input throughout their whole life.

One specific property of bilingual exposure is the existence of increased speech-input-variability, in particular in word pronunciation. Indeed, in bilingual environments, non-native speakers produce “accented” utterances and people living in bilingual environments are thus, on average, more frequently exposed to higher variability than people living in monolingual environments. This kind of variability in the daily speech input from very early on shapes a bilinguals’ speech perception system in a way that leads to more flexibility at the moment of word-identification by accepting accented – or mispronounced - utterances (Ramon-Casas, et al., 2009; Sebastián-Gallés, et al., 2005; Sebastián-Gallés, et al., 2006; Sebastián-Gallés, et al., 2009). However, it is unknown if this increased flexibility extends to other types of variability in the signal. The present research aims at investigating the consequences of life-long exposure to high-variability bilingual input on the mechanisms of word recognition and the encoding of indexical (speaker) variability.

2.1.1 Variability in spoken word recognition

Different studies have shown that rapid adaptation to new voices and dialects occurs (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Dahan, Drucker, & Scarborough, 2008; Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Maye, Aslin, & Tanenhaus, 2008; Mitterer & McQueen, 2009; Norris, McQueen, & Cutler, 2003) and that exposure to a high number of different exemplars (e.g. accented words) leads to better recognition of new words (Floccia, Goslin, Girard, & Konopczynski, 2006). These studies provide indirect support to the notion that increased variability in bilingual exposure may benefit their processing of highly varied speech, at least when compared with monolinguals. However, the processing of varied speech input comes with some cost. Certainly, the negative effects of variability on word recognition in monolingual populations have been attested in several works. These studies have indicated that variability in the form of different speakers, voice qualities and accents can negatively influence word-recognition in different tasks (Bradlow, et al., 1999; Church & Schacter, 1994; Goldinger, 1996; Nygaard, et al., 1995; Palmeri, et al., 1993; Schacter & Church, 1992). Different proposals have tried to explain how the system deals with this so called indexical variability that contains useful information but can act as noise at the speech perception level.

On the one hand, some models have suggested that abstract representations of words are stored in the mental lexicon and a normalized input, stripped of all indexical, irrelevant information, is matched to those abstract representations (abstractionist models: (Luce & Pisoni, 1998; Norris, 1994). On the other hand, it has been proposed that the mental lexicon is formed by veridical copies of each word we hear in our lives (exemplar models: (Goldinger, 1998; Pisoni, 1997). Substantial experimental evidence supports both accounts and recently hybrid models incorporating abstract and detailed episodic information have been put forward. Bowers and Kouider, (2003) proposed a model where only abstract, but not specific perceptual information is represented in the mental lexicon. Furthermore, this specific information, although retained in memory does not influence lexical decisions as it is not necessary for word recognition. McLennan and Luce, (2005) presented a time-course hypothesis, which states that depending on the time a subject has to give a response, slow-acting episodic traces or fast-acting abstract representations influence processing (see Orfanidou, et al., 2011 for a similar proposal). McLennan & Luce, (2005) used a lexical decision task and manipulated the speed with which participants gave responses by varying the similarity of non-words to real words. They found indexical effects only for difficult lexical discrimination (slow answers) and in a delayed response condition. Similarly, other studies have found indexical specificity effects only in conditions of deep effortful processing (Goldinger, 1996) and slow hard tasks (McLennan, Luce, & Charles-Luce, 2003). In general, these models assume that in the online processing of spoken words, indexical information does not play an immediate role and that for the most part word recognition processes are performed on abstract linguistic representations (McLennan & Luce, 2005).

There is a general consensus in that explicit memory tasks are likely to give raise to episodic effects, because they involve (effortful) slow responses. Indeed, using a recognition memory procedure, a variety of studies have observed that fine grained indexical detail influences the process of word recognition (Bradlow, et al., 1999; Church & Schacter, 1994; Goldinger, 1996; Nygaard, et al., 1995; Palmeri, et al., 1993; Schacter & Church, 1992). In the recognition memory paradigm participants have to decide whether a word is presented for the first time (“new”) or if it was already presented before (“old”), independently of the voice uttering the stimulus. The results generally show better performance in the case of repeated words when the voice uttering the word is the same than when the voice has changed. In case the stimulus was presented before in the same voice, both abstract- lexical information and episodic information coincide: the word was already presented and the voice was the same, so recognition is facilitated and performance is improved. However, in case of a change in voice, there is a conflict in the information provided at the abstract level (the word is “old”) and the information provided at the episodic level (the token is “new”). Because of this conflicting information participants slowdown their responses and respond “new” to “old” items more often.

2.1.2 Bilingual speech perception

As mentioned earlier, a bilingual environment poses additional challenges to the speech perception system. In addition to the dialectal variations in speech input that monolingual listeners are usually faced with, bilinguals have to deal with many mispronunciations of non-native speakers. It has been shown that a bilingual’s lexicon, in fact, contains accented entries as well as standard forms (Sebastián-Gallés, et al., 2005; Sebastián-Gallés, et al., 2006; Sebastián-Gallés, et al., 2009). Also if the two languages are present simultaneously in most situations in daily life chances are high that they are accidentally mixed and some words of language A might intrude into sentences uttered in language B. Although this might affect mainly selection of the appropriate word candidate and language, it also introduces another dimension of variability. Basically, a bilingual needs to be prepared for a broader spectrum of possible utterances (accented, mispronounced or even from the other language) at all times. More widespread activation would be a consequence and a mechanism for filtering and rapid selection of appropriate candidates would be necessary. Due to results on production (Gollan, et al., 2007; Roberts, et al., 2002) and reading (Ransdell & Fischler, 1987) it has been widely accepted that bilinguals are slower at lexical access, however at present we are not aware of any study showing slower lexical access during auditory speech perception for bilinguals. In fact in Pallier, et al., (2001) bilinguals and monolinguals performed an auditory lexical decision task and no reaction time differences were observed. This suggests that bilinguals have developed a mechanism to deal with the variable input in a fast and efficient manner.

Another source of additional variability in the process of word recognition in bilinguals comes from the non-selective activation of the bilinguals' lexicon. Several studies by Marian and colleagues (Marian, et al., 2008; Marian & Spivey, 2003; Marian, et al., 2003; Spivey & Marian, 1999) have shown that during auditory speech perception adult bilinguals access their two lexicons in parallel even in monolingual settings. For example Marian and Spivey, (2003) conducted an eye-tracking experiment looking at within-language and between language competitor fixations and have found similar results for both conditions, indicating simultaneous access to both lexicons.

Other studies have reported "spurious" activation of bilinguals' L2 lexicon due to lack of sensitivity to non-native phoneme contrasts. Pallier, et al., (2001) tested Spanish-Catalan bilinguals on a lexical decision task. Stimuli were minimal pairs including a Catalan-specific vowel contrast that Spanish natives found difficult to perceive. The results showed that Spanish dominant bilinguals treated those items as homophones resulting in repetition priming. Similarly, Sebastián-Gallés, et al., (2005) found that Spanish dominant bilinguals performed significantly worse on a lexical decision task involving non-words based on some Catalan-native contrasts. Interestingly, even natives who had been exposed to accented versions of Catalan had trouble rejecting some non-words that sounded the way Spanish-dominant bilinguals pronounced Catalan words (see also Sebastián-Gallés, et al., 2006; Sebastián-Gallés, et al., 2009 for studies using ERP measurements). Taken together, those studies show that bilinguals have to deal with more widespread activation in the lexicon as compared to monolinguals, therefore increasing the need of filtering out the relevant information from all the irrelevantly activated representations.

2.1.3 Processing of indexical information by bilinguals

In general, these studies report interference effects from one language to the other (slower response times or fixations on irrelevant distracters). However, dealing with these more challenging situations on a daily basis may also have beneficial effects for bilinguals when faced with high loads of variability in speech perception; for example when processing indexical variability such as accents or different speakers. Having practised to overcome the influence of several different sources of irrelevant variability on a daily basis may have lead to the development of mechanisms that are better able to deal with increased variability compared to monolingual speech perception mechanisms.

Therefore, if bilinguals are better at ignoring irrelevant information due to life-long practise with increased variability in the speech input, we expect differences in performance between bilinguals and monolinguals on a recognition memory task where stimuli are produced by different speakers. A recognition memory task requires participants to explicitly determine if a word is heard for the first time or if it is a repetition (irrespective of the uttering voice). This is a particularly effortful task and responses are quite slow. According to the models reviewed above (McLennan & Luce, 2005; Orfanidou, et al., 2011) slow responses are likely to be influenced by indexical/episodic information as well as by the information stored in the mental lexicon. In particular, we expect that bilinguals may be better at ignoring the irrelevant indexical (speaker) information and thus, show equivalent recognition memory for words repeated in the same voice and for words repeated by a different speaker. In contrast, monolinguals should show reduced (or no) recognition memory in the case of a voice change. These differences between monolinguals and bilinguals as a function of speaker change in the repetition of words should disappear when the task does not require complex, and slow, computations. As said, the influence of indexical information should diminish (or even disappear) in the case of easy/fast responses. We will test this hypothesis by comparing the effects of speaker change between monolinguals and bilinguals in a lexical decision task.

The bilingual society in the Barcelona area represents an ideal setting to test the present hypotheses. Catalan and Spanish (two Romance languages) are spoken by the majority of the population from a very early age on (97% of the population understands both languages and 85% speaks them; (Vila i Moreno, Vial, & Galindo, 2004). Importantly, Catalan and Spanish, although both Romance languages, differ in their phoneme inventory. Thus, some Catalan contrasts are hard to pronounce (and even perceive) for Spanish native speakers and chances to encounter mispronunciations are high. Consequently, living in Barcelona people have to manage two languages on a daily basis and are exposed to mispronunciations by

non-native speakers of either language (see Bosch & Ramon-Casas, 2011; Costa, et al., 2008 for descriptions of the bilingual environment).
In the present research, we compared Spanish monolinguals from the Madrid Metropolitan area and Spanish-Catalan bilinguals from the Barcelona Metropolitan area testing their capacity to recognize spoken words repeated by the same or a different speaker using a continuous recognition memory and a lexical decision task. In experiments 1 and 2 participants had to decide whether words were presented for the first time or as a repetition, regardless of the voice producing them. In experiment 3 we asked participants to perform a lexical decision task on stimuli produced by two different voices. We expect that the bilinguals' lifelong practise with increased variability in speech input will help them to ignore the irrelevant information conveyed by speaker changes. In accordance with hybrid models of word recognition, the differences between monolingual and bilingual participants should be larger in slow and effortful tasks (experiments 1 and 2) than in tasks involving fast responses (experiment 3).

2.2 Experiment 1

2.2.1 Background

Indexical variability such as speaker variability can disrupt priming effects on explicit recognition memory tasks (Church & Schacter, 1994; Palmeri, et al., 1993; Schacter & Church, 1992). For example, Palmeri, et al., (1993) conducted a series of continuous priming studies testing different numbers of speakers (2 speakers – 20 speakers) and also different lags between repeated items (2 – 64 intervening items between repetitions). They found consistently across all lags and number of speakers that a change in speaker produced higher reaction times and more errors than repetitions by the same speaker. Other studies using a blocked design also reported disrupted repetition priming for words produced by a different speaker at study and test (Church & Schacter, 1994; Schacter & Church, 1992).
Here, we conducted an explicit recognition memory experiment using stimuli spoken by a male and a female speaker. We used repetition priming effects to test the influence of indexical

variability on response times and performance. To investigate possible effects of practise with variability in the speech input we tested monolingual Spanish and bilingual Spanish-Catalan participants. We expected that monolinguals would show poorer performance in the recognition memory of items previously presented by a different speaker when compared with items repeated by the same speaker. In contrast, we expected similar performance of bilinguals when the repeated item had been produced by the same or a different speaker. Due to the lifelong exposure to many different exemplars and mispronunciations of words bilinguals should have less problems overcoming indexical variability in the form of speaker information than monolinguals.

2.2.2 Materials and methods

2.2.2.1 Participants

Eighty participants took part in the current study; forty Spanish-Catalan bilingual participants from the Universitat Pompeu Fabra, Barcelona, and 40 Spanish monolingual participants from the Universidad Autónoma in Madrid. Bilingual participants (age range 19 to 38 years; mean age: 23.95 years) had a high proficiency level in both Spanish and Catalan, which can be considered native languages. This was assessed using self-reported ratings on a 4 point scale. All bilingual participants were living in a bilingual environment at the time of testing and were using both languages on a daily basis. They had also all grown up in the same bilingual environment and had started learning both languages before school (see Appendix E). They had to show native performance and equivalent formal knowledge in both languages to enter University (as they had to pass the compulsory exam to enter any Spanish University that includes formal and informal knowledge of Spanish and Catalan). Of the 40 participants 28 were female and 12 male.
Monolingual participants had an age range of 18 to 23 years with an average age of 19.67 years. None of the monolingual participants was fluent in any other language than Spanish. In total 36 female and 4 male monolingual participants took part.
Although bilinguals and monolingual participants reported some knowledge of a foreign language (English) due to courses at school (see Appendix E), neither group was fluent in a foreign language.

All participants reported normal hearing and were paid for their collaboration

2.2.2.2 Stimuli

We recorded 144 Spanish words (critical items) in a sound attenuated booth by a male and a female speaker. The male speaker was from Madrid while the female speaker was a Spanish - Catalan bilingual from Barcelona. Stimuli were recorded in a sound-attenuated booth at a sampling rate of 44.1kHz on a PC using Cool Edit©. Background noise was filtered in Cool Edit© and amplitude of all stimuli was set at 65dB using PRAAT software (Boersma & Weenink, 2009). In addition, 50 words were recorded to serve as non-critical items. Of those, 30 were used to build up a memory load at the beginning of the experiment (so that a similar memory load is obtained throughout the whole experiment). The other 20 words were used as randomly distributed, non-repeated filler items (frequency range: 0.36 to 132.68 per million (all following word frequency values are per million) with a mean frequency of 20.762; Range of number of phonemes: 3 to 10 with a mean of 5.6; Range of number of syllables: 1 to 4 with a mean of 2.58; B-Pal: Davis & Perea, 2005; LEXESP: Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000). Of those 50 words, 25 were spoken by each speaker. We created four counterbalanced lists with the same order of items, just differing in speaker dimension on critical repeated items. For critical items, in terms of number of syllables (range from 1 to 5; mean: 2.6) and phonemes (range from 3 to 10; mean: 5.8) and word-frequency all lists were equal with a mean frequency of 35.56 ranging from 0.36 to 686.23 occurrences per million. An ANOVA on duration (speaker 1: 677ms; speaker 2: 533ms) found no differences between lists. All non-critical items were the same across lists. Overall the same number of primes was repeated by the same and the different speaker respectively.

2.2.2.3 Design and procedure

Participants listened to 338 words and had to decide for each of them if they had heard it for the first time or if it was a repetition (old-new decision). The procedure closely followed the continuous priming paradigm as described in Palmeri, et al. (1993). Participants were instructed to ignore the speaker and just answer if the particular word had appeared before or not. Items were repeated with a lag of 8-16 intervening stimuli, as this lag has been proven to produce priming effects, which could be influenced by speaker changes (Palmeri, et al., 1993). The interval between an answer and the onset of the next item was 1000ms. In case no answer was given the next item was played 1500ms after the previous stimulus. Participants were tested one at a time in a quiet room. The experiment was controlled by a PC using DMDX (Forster & Forster, 2003) software and stimuli were presented binaurally through Sennheiser 430 headphones. Control left and control right on the keyboard were used to respond; labels showing a "new" and an "old" were stuck to those keys to avoid left-right problems. Printed instructions indicated that subjects had to respond as fast as possible whether they had heard the stimulus for the first time or if it was a repetition, without committing too many errors. No further instructions were given as to how to press the buttons. The first 30 items were used to build-up a memory load and were not included in the analysis (see Palmeri, et al., 1993 for similar procedure). In Madrid procedures were the same except for the fact that participants were tested up to 8 at a time in separate cabins.

2.2.3 Results

Responses above 3000ms and below 200ms were excluded from the analysis. A by items analysis revealed at least 75% correct responses to all items, with no difference between groups ($t<1$), with the exception of 9 items with lower scores. All of those 9 items had similar response percentages between both groups ($t<1$).

We ran preliminary 2x2x3 ANOVAs with the between subject factor group (bilinguals, monolinguals) and the within subject factors speaker (speaker 1, speaker 2) and appearance (1^{st}, 2^{nd} same, 2^{nd} different). On neither reaction times nor error rates did we find a significant interaction of group by speaker nor any further interaction of those two factors with any other factor (all $F<1$). Expectedly, the factor speaker resulted in a main effect for reaction times ($F_{(1,78)}=957.313$; $p<0.0001$), as speaking rates differed between speakers (speaker 1: 677ms; speaker 2: 533ms). Error rates also displayed the same pattern ($F_{(1,78)}=221.857$; $p=0.009$). Because the factor speaker did not interact with any other factor, the main analyses were carried out without this variable.

2.2.3.1 Reaction times

A 2x3 ANOVA with group (bilinguals, monolinguals) as between subject factor and appearance (1st, 2nd same speaker, 2nd different speaker) as within subject factor was conducted on reaction times and error rates. Data on reaction times revealed a significant group effect ($F_{(1,78)}$=3.983; p=0.049) reflecting the fact that monolinguals (962ms) responded significantly faster than bilinguals (1007ms) overall. In addition, we found a significant main effect of appearance ($F_{(2,156)}$=423.309; p=.000; 1st : 1117ms; 2nd same: 920ms; 2nd different: 917ms). Subsequent post-hoc t-tests showed that a significant reduction in reaction times on the second appearance could be observed for same voice and different voice repetitions (1st - 2nd same: $t_{(79)}$=21.112; p<0.0001; 1st - 2nd different: $t_{(79)}$=22.649; p<0.0001), while there was no difference between the two on the second appearance (2nd same - 2nd different: t<1) (see Table 1).

Table 1 Reaction times (in ms) ± S.E. (standard error) of experiment 1 and 2 for bilinguals and monolinguals. First presentation collapsed across speakers and 2nd presentation in the same or different voice.

	Experiment 1 2 speaker		Experiment 2 8 speaker	
	Bilinguals	Monolinguals	Bilinguals	Monolinguals
1st (same and different collapsed)	1139 ± 20.9	1095 ± 20.9	1203 ± 27.8	1207 ± 27.8
2nd same	942 ± 14.8	898 ± 14.8	983 ± 20	975 ± 20
2nd different	941 ± 15	893 ± 15	990 ± 19.3	980 ± 19.3

2.2.3.2 Error rates

The same 2x3 ANOVA was run on the error rates. The factor group was marginally significant ($F_{(1,78)}$=3.217; p=0.077) and showed that monolinguals (13.63%) tended to make more errors overall than

bilinguals (11.75%). Errors did increase from first presentations to repetitions represented by a significant main effect of appearance ($F_{(2,78)}$=9.492; p<0.0001; see Figure 2). T-tests on appearance revealed that the first presentation differed significantly from both repetitions (1^{st} - 2^{nd} same: $t_{(79)}$=2.704; p=0.008; 1^{st} - 2^{nd} different: $t_{(79)}$=3.726; p<0.0001), while repetitions by the same speaker produced similar error rates as the repetition by a different speaker, reflected in a marginally significant difference (2^{nd} same vs 2^{nd} different: $t_{(79)}$=1.940; p=0.056). Hypothesis-driven analyses of each group separately were conducted. For bilinguals appearance resulted in a significant main effect ($F_{(2,39)}$=3.974; p=0.023; 1^{st} appearance: 9.45%; 2^{nd} same: 12.50%; 2^{nd} different: 13.30%). We ran t-tests and found that the first appearance produced less errors than both second appearances (1^{st} - 2^{nd} same: $t_{(39)}$=2.047; p=0.047; 1^{st} - 2^{nd} different: $t_{(39)}$=2.323; p=0.025). Repetitions (by same and different speaker) were equivalent between same voice and different voice repetitions (2^{nd} same - 2^{nd} different: (t<1). Thus, no speaker effect was detected for bilinguals (Figure 1).
Monolinguals also showed a significant main effect of appearance ($F_{(2,39)}$=5.710; p=0.005; 1^{st} appearance: 10.87%; 2^{nd} same: 13.89%; 2^{nd} different: 16.15%). While in contrast to bilinguals, the t-tests on appearance revealed differences between all three error rates. A marginally significant difference for 1^{st} vs 2^{nd} same speaker ($t_{(39)}$=1.778; p=0.083), a significant difference for 1^{st} vs 2^{nd} different speaker ($t_{(39)}$=2.902; p=0.006) and critically a difference between the repetition by the same speaker and the repetition by a different speaker ($t_{(39)}$=2.083; p=0.044).

2.2.3.3 d-prime

To further explore possible biases on error rates we calculated d-prime values for both conditions (same speaker and different speaker repetitions). Here we computed those items as first appearance which had actually served as primes for the repetition corresponding to the condition (same, different) in which it was repeated (not collapsing across first appearance). We ran a 2x2 ANOVA with the factors group (bilingual, monolingual) as between subject factor and condition (same, different) as within subject factor. While no main effects of either group or condition were seen, a significant interaction emerged ($F_{(1,78)}$=5.891; p=0.018).

Planned t-tests showed that bilinguals showed no difference in performance for same or different speaker conditions ($t_{(39)}$=1.050; p=0.300), while monolinguals performed significantly better on the same speaker condition ($t_{(39)}$=3.013; p=0.005) (see Figure 2). Comparing the conditions between groups, we found that bilinguals and monolinguals performed equally well on the same speaker condition (t<1), while bilinguals´ performance was significantly better than that of monolinguals in the different speaker condition ($t_{(78)}$=2.314; p=0.023). Thus, a bias-free measure reveals specificity effects for monolinguals, while bilinguals seem unaffected by a speaker change across measures.

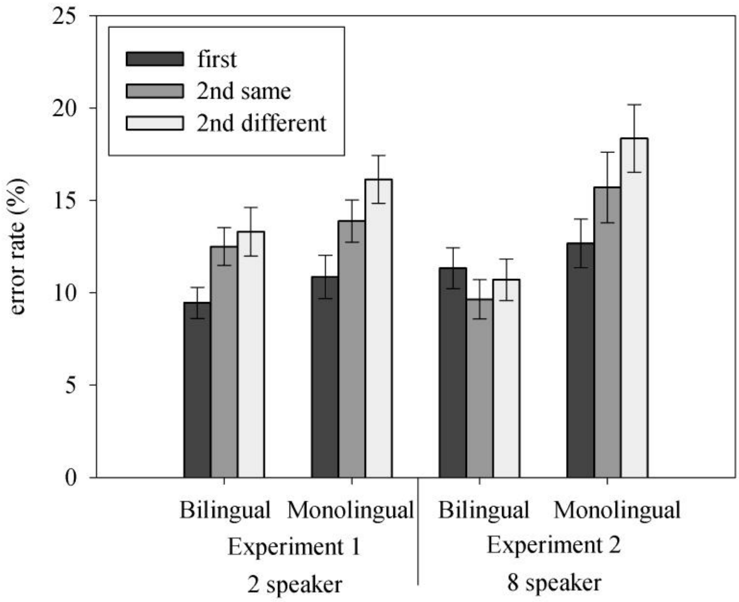

Figure 1 Mean error rate ± S.E. for the first appearance and the two repetitions in either the same voice or a different voice for both populations and experiments (experiment 1: 2 speakers; experiment 2: 8 speakers). In experiment 1 error rates tended to increase on repetitions and we can see that bilinguals committed the same amount of errors for both repetitions, irrespective of voice changes, while monolinguals produced more errors when the speaker changed from first presentation to repetition. In the 8 speaker task bilinguals tended to decrease the errors on repetitions, while monolinguals showed the same pattern as in the 2 speaker task.

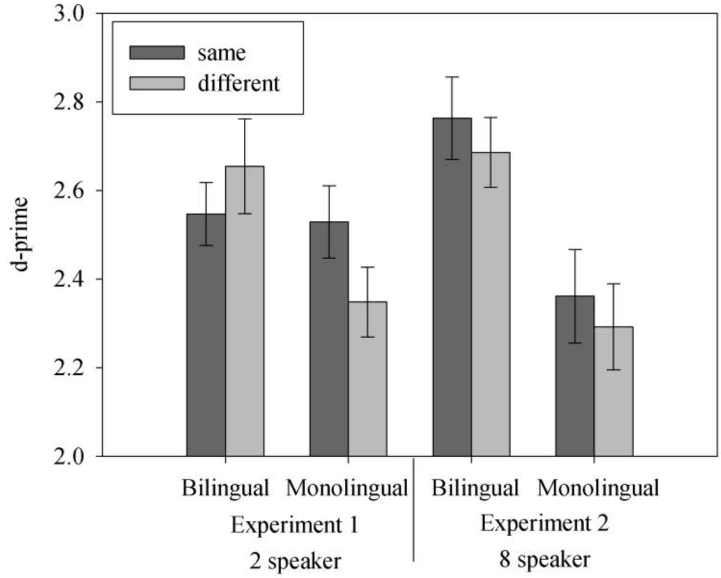

Figure 2 Mean d-prime ± S.E. results for the two conditions (same speaker and speaker change) and populations in both experiments. When processing 2 speakers bilinguals performed equally well for both conditions, while monolinguals´ performance was lower in the speaker change condition. In experiment 2 (8 speaker task) neither group showed a speaker effect, but only monolinguals decreased their overall performance while bilinguals showed a trend towards better performance.

2.2.4 Discussion

We found no effects of a change in speaker on reaction times in either group. On error rates only the monolingual group displayed an influence of the episodic speaker information, represented by a significantly increased error rate on the speaker change repetition compared to the repetition in the same voice. The d-prime measure revealed that bilinguals performed equally well in both conditions, irrespective of voice changes, and in addition, they tended to perform better than monolinguals especially on the speaker change condition. Monolinguals however, showed significantly worse performance when the speaker changed between presentations.

In the present study, we did not replicate studies finding an influence of indexical information on reaction time priming (Bradlow, et al., 1999; Church & Schacter, 1994; Goldinger, 1996; Nygaard, et al., 1995; Palmeri, et al., 1993; Schacter & Church, 1992), but our results are in in-line with some other publications reporting difficulties in obtaining specificity effects such as Luce & Lyons (1998), Gonzalez & McLennan (2007), Orfanidou, et al., (2011) and Orfanidou,et al., (2006). However, we did find specificity effects for monolinguals on performance, which suggests that bilinguals were less affected by the irrelevant speaker dimension and hence better at dealing with variability.
To further explore the influence of variability on speech perception, we wanted to test a situation with a comparably higher load of variability by increasing the number of speakers.

2.3 Experiment 2

2.3.1 Background

We ran a second study using the same task, but increasing the number of speakers and thereby indexical variability. We expected that monolinguals´ performance would decrease especially on error rates and d-primes compared to the two-speaker task. Bilinguals however, should not be significantly affected if they can successfully overcome interference from irrelevant information and thus, we expected similar error rates and d-prime values as in the previous task. Indeed, differences between bilinguals and monolinguals may increase with the number of speakers as bilinguals are particularly used to high levels of variability in the environment.

2.3.2 Materials and methods

2.3.2.1 Participants

Eighty participants took part in this study. Forty bilingual participants from the Universitat Pompeu Fabra in Barcelona, (age range: 18-43 years; mean 23.05 years; 20 male and 20 female participants) and 40 monolingual participants from the Universidad

Autónoma in Madrid (age range: 18-26 years; mean 19.12 years; 9 male and 31 female participants). All participants were from the same populations as in the previous study, but had not participated previously (see Appendix E for more detailed information on language background).

2.3.2.2 Stimuli

The same stimuli as in the previous study were recorded by six additional speakers (3 male and 3 female, 8 total with the male and female from the previous study) and filtered and processed to have the same volume and sound quality as the previous stimuli. Two male and two female speakers were native speakers of Castilian Spanish, while the other half were Spanish-Catalan bilingual speakers from the Barcelona Metropolitan area. Due to the number of speakers, it was not possible to counterbalance all items across all speakers and conditions. Thus, four lists were created with two sets of stimuli per speaker. Within in each set (8 items) half of the stimuli was repeated by the same speaker while the other half (4 items) was repeated by four other speakers (2 male, 2 female). Similarly, within the four lists half of the items were repeated by the same and the other half by different speakers. The same items appeared in all lists, so frequency (range from 0.36 to 689.8; mean: 33), number of syllables (range from 1 to 5; mean: 2.6) and phonemes (range from 3 to 10; mean: 5.8) were equal in all lists. An ANOVA looking at duration (speaker 1: 540ms; speaker 2: 679ms; speaker 3: 529ms; speaker 4: 745ms; speaker 5: 563ms; speaker 6: 605ms; speaker 7: 593ms; speaker 8: 658ms) found no difference between lists. To build up the memory load 32 items (4 by each speaker) were presented at the beginning (not included in analysis) and the additional 32 filler items were randomly placed across the task. Those 64 items were non-critical and thus not included in the analysis (frequency range: 0.36 to 278.04 with a mean frequency of 28.81; Range of number of phonemes: 3 to 10 with a mean of 5.7; Range of number of syllables: 1 to 4 with a mean of 2.5).

2.3.2.3 Design and procedure

Design and procedure were equivalent to the previous study except for the following: Participants listened to 288 words spoken by

eight different speakers (4 males, 4 females). Of those, 64 were non-critical items used to build-up the memory load and act as filler items.
Again, in Madrid procedures were the same except for the fact that participants were tested up to 8 at a time in separate cabins.

2.3.3 Results

We deleted all reaction times above 3000ms and below 200ms. We ran an analysis on items and found that all items were at least at 75% correct in both groups ($t<1$), except 11 items that were between 58% and 75% correct and had similar scores between groups ($t<1$). A 2x8x3 ANOVA on reaction times and error rates including group (bilingual, monolingual) as between subject factor and appearance (1st, 2nd same, 2nd different), and speaker (speaker 1-8) as within subject factors was conducted. On reaction times we found a main effect of speaker ($F_{(7,546)}=28.728$; $p<0.0001$) but neither an interaction of group by speaker nor further interactions of those two factors with any other factor were detected. Looking at errors we found a significant main effect of speaker ($F_{(1,78)}=2.883$; $p=0.006$) and an interaction of speaker by group ($F_{(7,546)}=2.056$; $p=0.047$) but no further interaction of those factors. As in this study 8 speakers were present, the interaction of speaker by group is complex in nature. Looking at the means revealed that bilinguals performed better than monolinguals for all speakers except speaker 2 (female, from Barcelona Metropolitan area), for which performance is equal for bilinguals and monolinguals.
A 2x3 ANOVA with group (bilinguals, monolinguals) as between subject factor and appearance (1st, 2nd same speaker, 2nd different speaker) as within subject factor was conducted on reaction times and error rates.

2.3.3.1 Reaction times

Reaction time results showed no effect of group ($F<1$) but a significant main effect of appearance ($F_{(2,156)}=310.466$; $p<0.0001$; 1st : 1205ms; 2nd same: 988ms; 2nd different: 985ms). Post-hoc t-tests revealed significantly smaller reaction times on the second appearance for both repetitions in the same and different voice (1st - 2nd same: $t_{(79)}=19.283$; $p<0.0001$; 1st - 2nd different: $t_{(79)}=17,910$;

$p<0.0001$), and no difference between same and different voice repetitions (2nd same - 2nd different: $t_{(79)}=1.204$; $p=0.232$) (see Table 1). We found no indication of a group by appearance interaction ($F<1$).

2.3.3.2 Error rates

On error rates we found a significant main effect of group ($F_{(1,78)}=13.484$; $p<0.0001$) showing that bilinguals (10.6%) committed significantly less errors across the whole experiment than monolinguals (15.6%). The factor appearance was not significant ($F_{(2,156)}=2.035$; $p=0.134$), while a significant interaction of appearance by group was found ($F_{(2,156)}=3.189$; $p=0.044$; see Figure 1). This interaction reflects the fact that error rates on repetitions increased for monolinguals, while they displayed the opposite trend for bilinguals. When running t-tests only for bilinguals, no differences were detected between all appearances (for 1st - 2nd same and 1st - 2nd different $t<0$; 2nd same - 2nd different $t_{(39)}=1.574$; $p=0.124$). For monolinguals however, errors increase marginally from first to second presentation in the same speaker condition (1st-2nd same: $t_{(39)}=1.778$; $p=0.083$) and from first to second presentation in the different voice we observed a significant increase (1st - 2nd different: $t_{(39)}=2.902$; $p=0.006$). Also, a significant difference between repetitions in same and different voice was present (2nd same - 2nd different: $t_{(39)}=2.083$; $p=0.044$) (Figure 1).

2.3.3.3 d-prime

We calculated d-prime scores and ran a 2x2 ANOVA with the factor group (bilingual, monolingual) as between subject factor and condition (same, different) as within subject factor. A highly significant main effect of group indicates a difference in performance between populations, with monolinguals scoring lower than bilinguals ($F_{(1,78)}=12.051$; $p=0.001$; bilinguals: 2.7; monolinguals 2.3). Neither an effect of condition nor an interaction were seen ($F<1$) (Figure 2).
On this recognition memory task with increased speaker variability, bilinguals still showed no specificity effects or any influence of speaker information, while monolinguals showed larger error

increases on the speaker change condition. Overall performance was better for bilinguals while being as fast as monolinguals.

2.3.3.4 Pooled Analysis

Based on the hypothesis that differences between bilingual and monolingual populations would increase with higher variability in the input, we wanted to compare the results from the two recognition memory studies presented above. Thus, we pooled data from both studies and ran a 2x2x3 ANOVA adding the between subject factor “number of speakers” (2 speakers, 8 speakers).

Reaction times

On reaction times we found a significant effect of number of speakers ($F_{(1,156)}$=14.694; p=0.000), but no effect of group ($F_{(1,156)}$=1.747; p=0.188). Overall reaction times were higher for experiment 2 using 8 speakers than for experiment 1 where 2 speakers had to be processed (2 speaker: 985ms; 8 speaker: 1056ms). A significant main effect of appearance ($F_{(2,312)}$=701.194; p<0.0001) was observed. T-tests showed that again there were no differences between repetitions by the same or a different voice (t<1), while both repetitions were faster than the first presentation (Table 1).

When considering monolinguals and bilinguals separately the same pattern of results was observed with a significant main effect of number of speakers (bilinguals: $F_{(2,156)}$=4.047; p=0.048; monolinguals: $F_{(2,156)}$=11.344; p=0.001) and a significant effect of appearance (bilinguals: $F_{(2,156)}$=294.202; p<0.0001 ; monolinguals: $F_{(2,156)}$=426.945; p<0.0001).

Error rates

The analysis on error rates showed no effect of number of speakers (F<1), but a significant effect of group ($F_{(1,156)}$=16.037; p<0.0001; bilinguals: 11.2%; monolinguals: 14.6%) and a marginally significant interaction of those factors ($F_{(1,156)}$=3.310; p=0.071), showing that bilinguals committed less errors on the 8 speaker task (2 speakers: 11.8%; 8 speakers: 10.6%), while monolinguals had

higher error rates when processing more speakers (2 speakers: 13.6%; 8 speakers: 15.6%).
We also found a significant effect of appearance ($F_{(2,312)}$=8.917; p<0.0001), with increased error rates on repetitions. Post-hoc t-tests showed differences not only for the first appearance compared to repetitions (1^{st} (11.1%) - 2^{nd} same (12.9%): $t_{(159)}$=1.940; p=0.054; 1^{st} - 2^{nd} different (14.6%): $t_{(159)}$=3.532; p=0.001), but also between same speaker and speaker change repetitions (2^{nd} same - 2^{nd} different: $t_{(159)}$=3.571; p<0.0001). Additionally, a marginally significant interaction of appearance with group ($F_{(2,312)}$=2.689; p=0.070) represented the fact that averaged across both experiments bilinguals showed no increase or decrease in error rates on repetitions (1^{st}: 10.4%; 2^{nd} same: 11.1%; 2^{nd} different: 12%; 1^{st} - 2^{nd} same t<1; 1^{st} - 2^{nd} different: $t_{(79)}$=1.263; p=0.210; 2^{nd} same - 2^{nd} different $t_{(79)}$=1.407; p=0.163) , while for monolinguals repetitions elicited increased errors (1^{st} (11.8%) - 2^{nd} same (14.8%): $t_{(79)}$=2.030; p=0.046; 1^{st} - 2^{nd} different (17.3%): $t_{(79)}$=3.589; p=0.001; 2^{nd} same - 2^{nd} different $t_{(79)}$=3.645; p<0.0001) especially on the speaker change condition (Figure 1).
We ran separate analyses for bilinguals and monolinguals and found no significant main effects for bilinguals. Only the interaction between appearance and number of speakers reached significance ($F_{(2,156)}$=3.125; p=0.047). Again, this represents the fact that bilinguals increased errors on repetitions in the two speaker task, while decreasing error rates were observed when eight speakers were processed.
For monolinguals a significant main effect of appearance was found ($F_{(2,156)}$=8.922; p<0.0001), but no other effects reached significance. T-tests confirmed that for both conditions error rates were significantly higher on repetitions (1^{st} (11.7%) - 2^{nd} same (14.8%): $t_{(79)}$=2.030; p=0.046; 1^{st} - 2^{nd} different (17.3%): $t_{(159)}$=3.589; p=0.001), while errors for the speaker change condition were also significantly higher than for the same speaker condition (2^{nd} same - 2^{nd} different: $t_{(79)}$=3.645; p<0.0001).

d-prime

The d-prime analysis with the factors group (bilingual, monolingual) and number of speakers (2, 8) as between subject factors and condition (same, different) as within subject factor revealed a highly significant main effect of group ($F_{(1,39)}$=12.939; p<0.0001), representing better performance of bilinguals across tasks. Separating groups led to no significant effects for bilinguals, showing similar performance across number of speakers and conditions. Monolinguals showed a main effect of condition ($F_{(1,39)}$=6.383; p=0.014), with worse performance for the speaker change condition across speaker variability levels (Figure 2).

2.3.4 Discussion

Considering the results of experiments 1 and 2 together, specificity effects only emerged for the monolingual population. While reaction times were unaffected by speaker changes, error rates and d-prime measures revealed that monolinguals showed larger error increases and an overall worse performance when the speaker changed between stimulus presentations. Thus, the monolingual data replicates previous research suggesting that surface detail of auditory input is retained and can influence performance in recognition memory.
The bilinguals however behaved in a different way. Not only was their performance not influenced by speaker information, but they also displayed a trend towards increased performance when speaker variability was higher.

2.4 Experiment 3

2.4.1 Background

The above discussed recognition memory studies are based on explicit memory components and require the participant to actively remember the items. Here it looks like bilinguals have a specific ability to recognize words without interference from irrelevant speaker variability. To test whether the above reported results can be replicated using a different task, we ran a similar priming study, but using a lexical decision task, which is based on implicit memory components.
In a lexical decision task, responses to words can be based on activation of abstract lexical entries and we expect no modulation of the repetition effects as a function of speaker change (Jackson & Morton, 1984; Luce & Pisoni, 1998; Norris, 1994). For non-words

however, predictions, and previous results are variable. Some studies have reported facilitatory repetition effects (e.g. Bowers, 1996; McKone & Dennis, 2000; Mimura, Verfaellie, & Milberg, 1997), while others failed to record any reaction time advantage for repeated non-words (Brown & Carr, 1993; Duchek & Neely, 1989; Ratcliff, Hockley, & McKoon, 1985). Repetition priming for non-words cannot be achieved through the lexicon and thus, priming effects may be based on episodic traces that have been formed on hearing the item for the first time, or alternatively activation of similar real words supports priming effects (Dorfman, 1994; Humphreys, Besner, & Quinlan, 1988). Orfanidou, et al., (2011) used a lexical decision task with two speakers and while they showed that word priming was speaker independent and thus providing evidence for abstract theories, they also found priming for non-words. They conclude that both abstract and episodic information influence speech perception in lexical decision tasks.
In the present study, we wanted to explore if those differences between monolingual and bilingual populations extend to an implicit memory task, which is based more on lexical access. It could be possible that rapid adaptation mechanisms allow bilinguals to overcome variable speech input, during explicit and implicit priming and maybe also independently of lexical representations. If this is the case, we expect bilinguals to perform better than monolinguals also on non-words.

2.4.2 Materials and methods

2.4.2.1 Participants

Eighty undergraduate students took part in the current study. Forty were bilinguals from the University of Barcelona (age range: 18-24 years; mean 20.56 years; 5 male and 35 female participants) and 40 were monolinguals from the University Autónoma in Madrid (age range: 17-30 years; mean 21.86 years; 15 male and 25 female participants). All participants were from the same populations as in the previous experiments, but had not participated before (see Appendix E for additional information).

2.4.2.2 Stimuli

Sixty Spanish words from the previous studies were used and 60 non-words were newly recorded in a sound attenuated booth by the same male speaker and female speaker as in Experiment 1. Pseudo-words were created by replacing one phoneme and can thus be considered word-like non-words that follow the rules of Spanish phonotactics. Stimuli were recorded at a sampling rate of 44.1kHz on a PC using Cool Edit©. Background noise was filtered in Cool Edit© and amplitude of all stimuli was set at 65dB using PRAAT software. Of the 60 words (same for the non-words) half were spoken by speaker 1 and the other half by speaker 2. For each speaker 20 words (and 20 non-words) were critical repeated stimuli and 20 were non-critical fillers. Four experimental lists were created, which were identical except for the speaker on repeated items (see Table 1). Thus, each list contained equal numbers of primes that matched targets (40) and primes that mismatched targets (40) on speaker dimension. All lists contained the same items, thus they were all equal in terms of word-frequency (range from 0.36 to 689.8; mean: 38.9 occurrences per million), number of syllables (range from 1 to 5; mean: 2.6) and number of phonemes (range from 3 to 10; mean: 5.8). An ANOVA was conducted on duration (speaker 1: 673ms; speaker 2: 546ms), which did not result in significant differences between lists.

2.4.2.3 Design and procedure

A continuous priming experiment was conducted using an auditory lexical decision task. Participants heard a series of 240 stimuli spoken by two speakers (male and female), these included 80 non-critical filler items (40 words and 40 non-words) and for each speaker 20 repeated critical words (20 first appearances and 20 second appearances) and 20 repeated non-words. The experiment consisted of one block of 240 items with 8-16 intervening items between first (prime) and second (target) appearances (similar procedure as in previous experiments). When participants heard an item they had to decide as quickly and accurately as possible, if the stimulus was a word or not. The time between an answer and the onset of the next item was 1000ms. In case no answer was given the next item was played 1500ms after the previous stimulus.

Participants were tested four at a time (one participant/list) in separate cabins. The experiment was controlled by a PC using EXPE software (Pallier, Dupoux, & Jeannin, 1997) and stimuli were presented binaurally through Sennheiser 430 headphones. A response box with two buttons was used with the right button indicating a non-word and the left button indicating a word. Participants were instructed to press the buttons using two fingers of their dominant hand.

In Madrid, testing-conditions were the same except for the fact that participants were tested up to eight at a time in separate cabins using the keyboard of the computer (control right and left marked with a "1" and "2") to give responses.

2.4.3 Results

Responses with a reaction time of 200ms or lower and trials with no response (longer than 3000ms) were discarded. Less than 1% of the trials had to be rejected due to this criterion. It was found that all items were responded to correctly by at least 75% of participants in both populations. None of the items differed significantly between groups ($t<1$) in terms of % correct responses.

First, we ran a 2x2x3 ANOVA including group (bilingual, monolingual) as between subject factor and speaker (speaker 1, speaker 2) and appearance (1st, 2nd same, 2nd different) as within subject factors. On reaction times we found a significant effect of speaker due to differences in speaking rate ($F_{(1,78)}=57.133$; $p<0.0001$) and a speaker by group interaction ($F_{(1,78)}=8.145$; $p=0.006$) but no further interactions of those factors with any other factor. This interaction showed that the faster group (bilinguals) had a difference in reaction times between the two speakers while monolinguals had equal reaction times across speakers. Due to the fact that mean duration of stimuli for speaker 1 (673ms) were longer than those of speaker 2 (546ms), it is not surprising that the faster participants might be affected more by a difference between durations. On error rates neither the main effect of speaker ($F_{(1,78)}=2.332$; $p=0.131$) nor the speaker by group interaction ($F_{(1,78)}=2.362$; $p=0.128$) reached significance. No further interactions of group by speaker with any other factor were found showing that both speakers were equally well processed.

2.4.3.1 Reaction times

We ran a 2x2x3 ANOVA with group (bilingual, monolingual) as between subject factor and lexical status (word, non-word) and appearance (1^{st}, 2^{nd} same, 2^{nd} different) as within subject factors.
Confirming our hypothesis that processing of words and non-words would be different for bilinguals and monolinguals, we found a significant interaction of group by lexical status ($F_{(1,78)}=91.603$; $p<0.0001$), representing the fact that the reaction time increase from words to non-words was much larger for monolinguals (250ms) than for bilinguals (126ms). Thus, we ran separate analysis on words and non-words (Figure 3).

Words

Running a 2x3 ANOVA with group (bilingual, monolingual) as between subject factor and appearance (1^{st}, 2^{nd} same, 2^{nd} different) as within subject factors, we found a significant effect of group ($F_{(1,78)}=14.831$; $p<0.0001$), showing that bilinguals (768ms) responded faster than monolinguals (843ms). Significant repetition priming for words was found across both groups indicated by a significant main effect of appearance ($F_{(2,156)}=23.211$; $p<0.0001$; a t-test including both groups confirmed that across groups no difference between same and different speaker repetitions was seen $t_{(79)}=1.072$; $p=0.287$). This is indicating that both groups showed priming irrespective of the voice.

Non-words

The same 2x3 ANOVA as for words was run on non-words. We found a significant group effect ($F_{(1,78)}=87.178$; $p<0.0001$; bilinguals: 894ms; monolinguals: 1094ms), a significant main effect of appearance ($F_{(2,156)}=10.153$; $p<0.0001$), and a marginal interaction of the two factors ($F_{(2,156)}=2.403$; $p=0.094$).
We conducted separate analysis on non-words for each group and found that only bilinguals displayed repetition priming for non-words represented by a significant main effect of appearance ($F_{(2,78)}=13.279$; $p<0.0001$), which was absent for monolinguals ($F_{(2,78)}=1.868$; $p=0.161$). T-tests for bilinguals showed again a significant difference for first presentation compared to both

repetitions and no difference between repetitions in the same or different voice (see Table 2). Monolinguals however, showed only a marginally significant difference between first presentation and repetition in the same voice ($t_{(39)}$=1.851; p=0.072). Thus, a trend towards a speaker effect emerged only for monolinguals.

Table 2 Priming effect ± S.E. in ms (mean difference) and t-test values for significant and marginal results on reaction times for non-words.

Non-words	$1^{st} – 2^{nd}$ same	$1^{st} – 2^{nd}$ different	2^{nd} same – 2^{nd} different
Bilinguals[(a)]	37 ± 8.1 4.489; ***	34 ± 7.1 4.807; ***	2 ± 8.5 n.s.
Monolinguals[(a)]	19 ± 10.3 1.851; +	6 ± 8.8 n.s.	13 ± 10.8 n.s.
Both groups[(b)]	32 ± 4.1 7.669; ***	25 ± 3.9 6.384; ***	7 ± 4.8 n.s.

[(a)] t(39); [(b)] t(79); *** p<0.0001; + p=0.072; n.s.=not significant

2.4.3.2 Error rates

The same 2x2x3 ANOVA was conducted on error rates. Again, we observed a significant interaction of group by lexical status ($F_{(1,78)}$=4.602; p=0.035). Similarly as for reaction times, the error rate increase from words to non-words was significantly larger for monolinguals (4.1%) than bilinguals (1.9%). Again we ran separate analysis for words and non-words. Error rates were very low and analyses should be considered with caution (Figure 4).

Words

We found a marginally significant difference between groups ($F_{(1,78)}$=3.443; p=0.067; bilinguals: 1.7%; monolingual: 2.6%), and a significant effect of appearance ($F_{(2,156)}$=3.854; p=0.023). (T-tests revealed a significant error decrease only between first presentation and the repetition in the speaker change condition (for group analysis: $t_{(79)}$=3.182; p=0.002; and monolinguals only: $t_{(79)}$=3.819; p<0.0001; for monolinguals also between 2^{nd} same and 2^{nd} diff: $t_{(79)}$=1.912; p=0.063).

Non-words

Considering only non-words we found a significant difference between groups ($F_{(1,78)}$=9.578; p=0.003), but no effect of appearance and no interaction (all F<1). Looking at the means revealed that bilinguals (3.6%) committed significantly fewer errors on non-words than monolinguals (6.6%).
Taken together, performance for words was similar between groups, while for non-words monolinguals scored lower and responded slower than bilinguals.

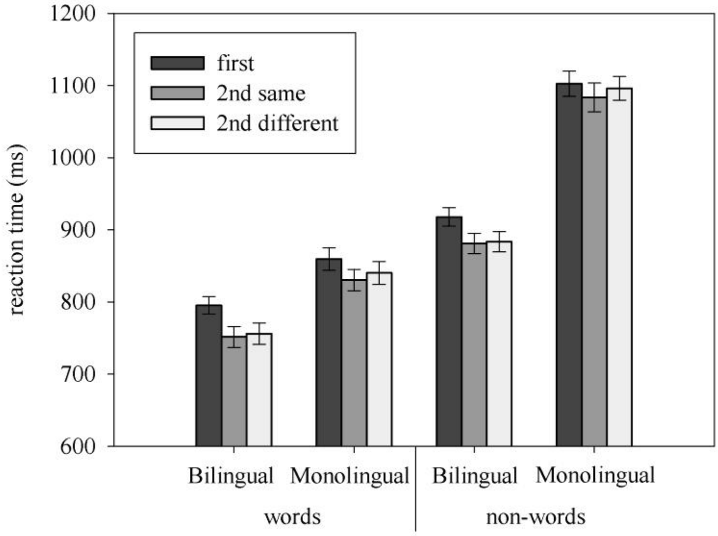

Figure 3 Mean reaction times ± S.E. for words and non-words for both groups. All participants responded faster and showed significant priming for words, while only bilinguals also showed priming for non-words. Also, the difference in reaction times between words and non-words was much bigger for monolinguals.

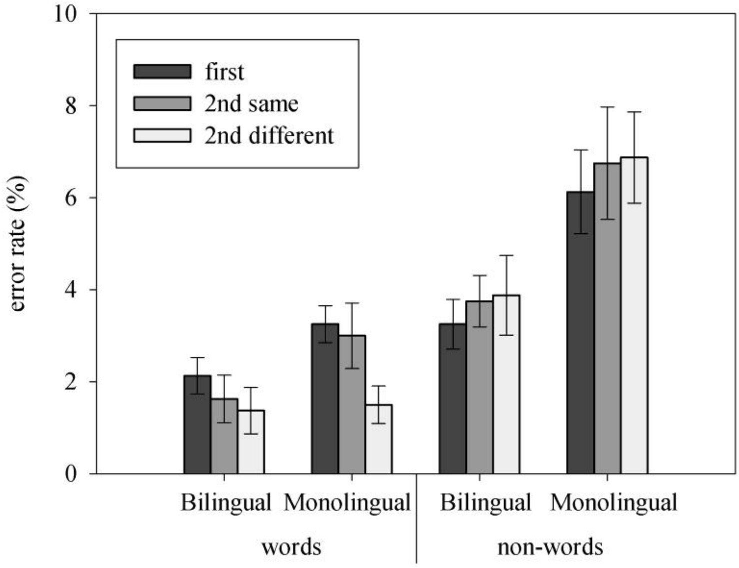

Figure 4 Mean error rates ± S.E. for words and non-words committed by bilinguals and monolinguals. No priming effects were observed, except for words in monolinguals and interestingly, only in the speaker change condition. Again, error increase from words to non-words was much higher for monolinguals than bilinguals.

2.4.4 Discussion

In neither population nor measure we found clear significant differences of same speaker versus different speaker repetitions, indicating that no specificity effects were recorded. In other words, reaction times and error rates were not affected by a change in speaker for either population. Interestingly, bilinguals responded faster and made fewer errors than monolinguals, while also being less influenced by the lexical status of the items, represented by the decreased errors and reaction times on non-words compared to monolinguals. Here, in contrast to the explicit recognition memory, where bilinguals were slower than monolinguals, bilinguals not only committed fewer errors but were also faster. Usually bilinguals were found to be slower on tasks implying lexical access (for speech production see Gollan, et al., 2007; Roberts, et al., 2002). However, as noted earlier no difference in reaction times was found for an auditory lexical decision task comparing bilinguals and monolinguals (Pallier, et al., 2001). One tentative explanation is that speaker variability was at the basis of the difference found in the present study. As shown bilinguals seem to be less affected by speaker variability and hence, when introducing variability in a lexical decision task a reaction time advantage for bilinguals emerged. Priming for words was observed in both populations and was not affected by speaker changes, which was expected due to the rapid responses that could be based on abstract lexical entries. Importantly, we found no priming for non-words in monolinguals, while bilinguals profited from repetition irrespective of the speaker. Thus, as we found that bilinguals performed better and faster especially on non-words, it suggests that rapid mechanisms to overcome variability in the input are at the base of our findings.

2.5 General discussion

Across all three experiments we found that bilinguals, in contrast to monolinguals, were not negatively influenced by indexical speaker information. In experiment 1, using an explicit recognition memory task and two speakers the bilinguals' performance was equivalent for same speaker and speaker change conditions, while monolinguals scored significantly lower when faced with items repeated in a different voice. Interestingly, when increasing variability from two to eight speakers (experiment 2), bilinguals exhibited a trend towards an overall better performance. This trend was mainly driven by the same speaker condition, on which performance clearly improved while different speaker performance stayed constant across experiment 1 and 2. Those results indicate that bilinguals are particularly able to deal with a high-variability environment. Monolinguals on the other hand, showed a trend toward an overall lower performance when the number of speakers was increased from two to eight. Considering d-prime scores, it seems that monolinguals decreased performance mainly on the same speaker condition, while performance on the speaker change condition remained about equal between experiments. In contrast to bilinguals, many speakers in the environment seem to dilute specific information for monolinguals, resulting in the loss of the advantage of coinciding information in the case of same speaker (same token) repetition. This is suggesting that in case of high variability-input monolingual participants were not capable of using speaker information, while for bilinguals same speaker information has become even more salient when processing many speakers.

In experiment 3, using a lexical decision task, we found that bilinguals were generally faster and committed fewer errors. This advantage was significantly higher when responding to non-words compared to words. Additionally, bilinguals showed repetition priming for non-words, which suggests that they can rapidly profit from episodic information while at the same time overcoming variability from speaker information. As discussed previously, both populations should not show speaker effects on word priming, as responses are rather quick and easy and can be mediated by the lexicon independently of episodic speaker information. As for non-words, interpretation of the results is more complicated. As mentioned before, some studies have found repetition priming for non-words (e.g. Bowers, 1996; McKone & Dennis, 2000; Mimura, et al., 1997; Orfanidou, et al., 2011; Orfanidou, et al., 2006) while other studies have not (Brown & Carr, 1993; Duchek & Neely, 1989; Pallier, et al., 2001; Ratcliff, et al., 1985). As non-words usually share portions with real words, it has been suggested that priming for pronounceable legal non-words might be based on the activation of those real words (Dorfman, 1994; Humphreys, et al., 1988). In light of a lexical decision task, this might introduce a bias towards incorrect responses. In the present study, only bilinguals showed priming for non-words, while at the same time committing significantly fewer erroneous responses than monolinguals. This suggests that practise in dealing with increased lexical activations during bilingual speech perception might have given bilinguals the ability to successfully use lexical activation of words to prime correct rejections of similar sounding non-words. In other words, a mechanism sorting through activated traces and selecting the appropriate candidate or response might be at the base of non-word priming and overall better performance in bilinguals. Also, as noted earlier, overall better and faster performance may be due to the speaker variability, which seems to selectively affect monolinguals more than bilinguals.

Considering together the results from all three studies, we can conclude that the environment in which bilinguals have acquired and used their languages has shaped their speech perception system to accommodate higher levels of variability in the signal. In neither experiment were bilinguals affected by variability from speaker information, while monolinguals showed lower performance when faced with speaker changes between repetitions. Also, bilinguals´ performance on non-words suggests a rapid mechanism to overcome negative effects of variability and at the same time retaining useful information (speaker independent representation and thus priming of non-words).

The larger number of activated entries (co-activated L2 entries and accented or mispronounced versions) in the bilingual lexicon during speech perception may be a potential origin of their better capacity to ignore irrelevant information in the speech signal, as they are used to extracting relevant from irrelevant activations. As said in the introduction, bilinguals are usually exposed to higher variability because in bilingual environments it is common that many speakers produce one of the languages with a foreign accent. Therefore, they are exposed to both native and non-native utterances. Among others, Sebastián-Gallés, et al., (2009) showed that the bilingual lexicon includes native and accented words. Thus, through intensive practice in managing variable speaker input, bilinguals seem to have developed a capacity to overcome interference from indexical information in speech perception. This may be achieved by a rapid selection mechanism that helps to identify and select the appropriate word candidate from the different activated entries. While bilinguals need to employ such a mechanism on a daily basis, for monolinguals it may only be necessary when encountering a strong dialect or foreign accent.

The notion that variability in the auditory speech input has consequences for bilingual speech perception is also supported by data from infant studies. Bilingual infants seem to have adopted strategies to deal with the increased load of learning two languages. When faced with minimal differences between tokens or minimal item pairs, they tend to ignore those differences until a later age than monolinguals (Albareda-Castellot, et al., 2011; Fennell, Byers-Heinlein, & Werker, 2007; Fennell & Werker, 2003; Ramon-Casas, et al., 2009). As shown by Bosch & Ramon-Casas, (2011), non-native bilingual speakers produce variable pronunciations (mispronunciations) and exposure to many different pronunciations might lead to insensitivity to small differences in phonology for infants, while possibly helping adults to understand many different speakers (and accents/dialects).

In the present study, we provided evidence that only monolinguals were affected negatively by a change in speaker between first presentation and repetition. In addition, the fact that increasing the number of speakers and thereby variability, selectively increased bilingual performance while decreasing monolingual performance, also supports the hypothesis that bilinguals are especially good at solving tasks involving a highly variable environment.

In the preceding we have entertained the hypothesis that the origin of the differences observed between monolinguals and bilinguals in our studies can be found in their better capacity to deal with variable input. However, an alternative hypothesis is also possible. There is a bulk of literature showing that bilinguals outperform monolinguals in several tasks, related to the functioning of cognitive control. The origin of this bilingual advantage is supposed to be the outcome of cognitive adaptations to cope with the word selection problem in speech production. Several studies have shown that the bilinguals' lexicons are accessed in parallel both when they recognize a word and when they have to produce it (Colomé, 2001; Costa, Caramazza, & Sebastian-Galles, 2000; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Marian & Spivey, 2003; Spivey & Marian, 1999 among others). While parallel activation of both bilinguals' lexicons may just increase the processing load in word recognition (as the sum of both lexicons includes more words), the situation is quite different in word production. To be able to lexicalize words in the intended language while preventing interference from the non-intended language, bilinguals need to involve some sort of language control mechanism. Although there is an intense debate in the field concerning the exact nature of this mechanism, there is agreement that some form of inhibition is involved (Costa, 2005; Costa, Miozzo, & Caramazza, 1999; Finkbeiner, Almeida, Janssen, & Caramazza, 2006; Green, 1998; Kroll, Bobb, & Wodniecka, 2006; La Heij, 2005). The continuous practice of these mechanisms would lead to a cognitive advantage in situations where conflicting information is present. Important to our research interests, as it will be seen further below, these advantages are present only when interference suppression is needed, while no advantage for response inhibition has been reported (Bialystok, Craik, & Ryan, 2006). Interference suppression is the concept of ignoring or suppressing misleading information that would lead to a false response, while response inhibition refers to the ability to refrain from making a false response in case of a salient cue (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Luk, Anderson, Craik, Grady, & Bialystok, 2010). In addition, advantages on conditions with high levels of incongruent trials indicate better monitoring capacities for bilinguals (Costa, et al., 2009). In light of those studies, an alternative explanation for our findings may be found in the better ability of bilinguals in resolving response conflict. In the recognition memory studies, a word repeated by a different speaker contains conflicting information, as it is an “old” item (same word) but a new token (by a different speaker). If speaker information is stored with episodic traces, being better at overcoming speaker information could be interpreted as better conflict resolution abilities. The fact that increasing variability in the environment (experiment 2) leads to a tendency of bilinguals to perform better, could be compared to results indicating lager differences between monolinguals and bilinguals when the level of incongruent trials is especially high (Costa, et al., 2009). In the lexical decision task where participants have to respond if stimuli are words or not, a conflict might be present between the familiarity of a repeated non-word and the “NO” response, assuming that performing lexical decision is based on familiarity. The overall better (and faster on lexical decision) performance of bilinguals could be interpreted as generally better monitoring capacities. However, as the task did not require tracking the specific voice uttering a stimulus this interpretation does not necessarily hold.

Although, our results in the recognition memory experiments may be partially explained by generally better abilities in cognitive control of bilinguals, the present results point in the direction that the origin of those effects may not solely lie in the word-selection problem in word production. Indeed, we propose that the influence of increased variability during speech perception plays a fundamental role. Even though no additional mechanisms may be required to select the appropriate word candidates in case the phonological information provided in the signal is clear, the increased number of activated lexical entries and the more detailed information about possible pronunciations of a given word represent an increased processing load. That bilinguals are dealing with this increased load with the same speed as monolinguals (whose processing load is smaller), suggests that adaptive strategies have shaped bilingual speech perception. How far those adaptations have also influenced other more general abilities and thus, also play a role in previous findings of bilingual advantages, remains to be explored in future studies. Further evidence that some differences between bilinguals and monolinguals originate from auditory speech perception is provided by a study by Kovács and Mehler, (2009) showing that seven-month-old bilingual infants show enhanced cognitive control mechanisms when compared to monolingual infants, well before they start to produce their first words.

Taken together, the two recognition memory experiments indicate that life-long practise with variable speech input, such as accented speech, and the resulting increase in lexical activations have lead to improved performance of bilinguals on tasks that vary the speaker information. Here bilinguals were not affected by irrelevant speaker information. In our case, when faced with many speakers (highly-variable environment), bilinguals tended to increase their performance while monolinguals did worse.

Interestingly, no specificity effects were observed on reaction times in either group or experiment. This might be due to the number of stimuli in our study compared to previous investigations and is in-line with other studies using a similar design which did not report specificity effects on reaction times (González & McLennan, 2007; Luce & Lyons, 1998; Orfanidou, et al., 2011; Orfanidou, et al., 2006).

Although the present research was not designed to test the question of how abstract and episodic information interact during auditory speech perception, we consider that the present results fit with the predictions of a hybrid model that integrates both processes depending on the task circumstances and requirements. For the monolingual group explicit priming processes (in the recognition memory experiments) relied more on episodic information, represented by speaker specificity effects, while during implicit priming (lexical decision), mainly abstract information was used, represented by the lack of priming for non-words and the absence of speaker effects. In the bilingual group such a clear separation was not possible. In the explicit recognition memory task bilinguals seem to have overcome irrelevant, variable information from episodic traces as shown by the lack of speaker effects. In the implicit lexical decision task priming for non-words suggests the use of episodic memory traces, while the absence of specificity effects again shows that influence from irrelevant variable information was overcome.

In conclusion, with regard to auditory speech perception, we found evidence for a hybrid model of abstract and episodic information, which interact differently depending on the task, and whether participants were bilingual or monolingual. In detail, we found that the episodic information influenced more explicit priming, while abstract information seems to be at the base of implicit priming. Also, we found that compared to monolinguals, bilinguals are selectively inhibiting conflicting indexical information when it would interfere with the task, while being able to use episodic information efficiently when it helps to solve the task.

Importantly, we have found evidence in three studies showing that early and highly skilled Catalan-Spanish bilinguals, who have been exposed to an environment with increased loads of variability in the speech input, can overcome interference from irrelevant variability in the form of speaker information. No indexical specificity effects were recorded and overall better (and faster on lexical decision) performance on all tasks suggests that bilinguals have developed strategies to rapidly deal with variability in speech input. Additionally, better and faster performance as well as speaker independent priming when processing non-words, indicates the involvement of a rapid on-line mechanism. General advantages on conflict resolution and monitoring of ongoing task requirements might also play an important role in the effects observed in the present study.

The exact nature of the mechanism proposed in the present study needs to be explored in more detail in future studies. It could be especially interesting to investigate the developmental path of such a mechanism by looking at young bilingual language learners and test their capacity to manage variable speech input. This could shed more light on the question of how much exposure to the variable environment is necessary for the emergence of the effects observed in the present study and maybe even entangle effects from speech production and speech perception.

3. RECOGNTION OF DEPICTED MISPRONOUNCED WORDS BY BILINGUAL AND MONOLINGUAL INFANTS

3.1 Introduction

The previously presented results on adult bilingual speech perception, point towards the development of mechanisms to effectively deal with increased variability in the environment. Now we can ask the question of when certain adaptations to the speech perception system are developed. It is possible that many years of exposure to a bilingual environment are necessary to adapt the system, while it is also possible that the basis already evolves during early language development. Previous research points towards differences between monolingual and bilingual speech perception as early as 4 months of age (Bosch & Sebastián-Gallés, 1997). As discussed in the general introduction, perceptual tuning to the native language occurs in the first months of life and being exposed to two languages requires different processes already at this stage. As shown in several studies, bilinguals seemed to not discriminate certain close contrasts that were part of one or both of their two languages (Bosch & Sebastián-Gallés, 2003; Sebastián-Gallés & Bosch, 2009). However, when using a different task it was found that bilingual infants did indeed discriminate the same contrast (Albareda-Castellot, et al., 2011). Similarly, in a word recognition task bilingual toddlers seemed to be insensitive to some mispronunciations (Ramon-Casas, et al., 2009). Taken together, these results suggest that bilingual infants only pay attention to changes between certain phonemes, in the cited studies vowels, if they are of relevance. Increased variability linked to the existence of two languages in the environment may have lead to the development of early adaptation mechanisms: it may be useful to pay less attention to variable structures as long as they do not carry crucial information, to reduce processing loads.

In the present study, we want to test bilingual and monolingual infants on a word-identification task involving variable pronunciations in vowels, to explore whether at an age of 18 - 20 months bilingual infants display adaptive strategies to overcome variable speech input.

In the last years a number of studies have explored processing of accents and mispronunciation in infants, however all those studies have looked at monolingual populations. In an early experiment, Nathan, Wells, & Donlan, (1998) conducted a word identification task with 4 and 7 year old children from London testing them on their own and a Glaswegian accent. It was found that word comprehension was disrupted in both age groups when listening to the Glaswegian accent. More recently, Best, et al., (2009) compared 15 and 19 month old toddlers on a familiarity preference task using their native (Connecticut English) and a non-native dialect (Jamaican Mesolect). Recognition of words was measured based on infants´ preference to listen to familiar words over unfamiliar ones. Hence, if they recognise familiar words produced with a foreign accent they should also show this familiarity preference when listening to the Jamaican accent. At 15 months of age familiarity preference was only observed for the native but not for the non-native dialect. However, at 19 months infants showed a preference for familiar words in both dialects. Those results indicated that at around 19 months of age infants had begun to discover phonological constancy, meaning that different pronunciations of the same word are possible. However, using a different task, where infants were presented with two pictures of familiar items and were instructed to look at the target item in either the native or the Jamaican dialect, no recognition of words pronounced in the Jamaican dialect was seen even in 18-20 month old infants (Mulak, Best, Irwin, & Tyler, 2008).

In a slightly different approach Schmale & Seidl (2009) tested 9 and 13 month old infants on a familiarisation task. Different speakers of English with either a native or a foreign (Spanish) accent were tested. At 9 months of age Spanish accented words were only recognised when previously familiarised with the same Spanish accented speaker. At 13 months infants were able to recognise words even when produced by two different Spanish accented speakers.

Studies on adult listeners, using a systematic shift in vowels, have shown that participants were able to rapidly adapt to the shift and generalise to new items (Kraljic & Samuel, 2005; Maye, et al., 2008; Norris, et al., 2003). Similarly, White & Aslin, (2011) systematically shifted the vowel [a] to [æ] and presented 18 – 20 month old infants with either the standard or the shifted pronunciation while showing pictures of the corresponding items. In a following test phase two objects were displayed and infants´ preferential looking to target and distracter objects was measured both with standard and shifted pronunciations. When previously exposed to the shifted pronunciation infants were able to relate this mispronunciation to the appropriate item, while the control group (exposed only to standard pronunciations) showed a trend towards recognition of mispronounced items only at the end of the test phase. This result showed that short exposure to a systematic shift can facilitate recognition of shifted pronunciations also in young learners.

Taken together those studies suggest that depending on the task, infants as young as 9 months can recognise items with accented pronunciation, when previously exposed, while at around 19 months of age toddlers seem to have started to adapt to unknown dialects.

As said, the world in which a bilingual infant grows up, however, presents some relevant differences when compared to the monolingual environment. Not only do bilinguals have to deal with two sets of grammar, phonology and vocabulary, but also the strategies employed by monolingual infants are not necessarily optimal for bilinguals.

When presented with pictures of a familiar and an unfamiliar item while listening to a new word, monolingual infants tend to look more to the unfamiliar object (applying the so-called "mutual exclusivity" heuristic). This has been interpreted as evidence that infants consider each object to have only one label (Markman, 1989, 1992; Merriman, 1991; Merriman & Bowman, 1989). Studies on infants learning more than one language however, have demonstrated that those infants rely significantly less on this word learning strategy than monolinguals (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Byers-Heinlein & Werker, 2009). Considering that in a bilingual environment two labels exist for each object it would not be useful for bilinguals to employ the same strategies as monolinguals.

Taking into account that bilingual infants are also exposed to many mispronunciations, it is possible that bilingual infants are better able to identify familiar objects when they are labelled with a new word that is only minimally different from the known label. Furthermore, in the case of bilingual environment where etymologically close languages, such as the case of Catalan and Spanish, many words are cognates, especially in a young infants´ vocabulary (Águila, Ramon-Casas, & Bosch, 2007; Pons & Sebastián-Gallés,

unpublished data), which means they share a great part of their phonemes; especially consonants are staying relatively constant across the two languages. Considering the increased processing load bilingual learners a faced with, paying more attention to common structures may be a useful mechanism to deal with the complex environment. Indirect evidence supporting this claim is provided by the fact that adult bilinguals process cognate items faster and more accurately than non-cognates (Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Marian & Spivey, 2003; Van Hell & Dijkstra, 2002), suggesting that common structures facilitate processing.
In the present study, we want to compare monolingual Spanish learners and bilingual Spanish-Catalan learners on a word recognition task using standard and systematically shifted pronunciations. We expect bilingual infants to accept mispronunciations without previous exposure, due to the fact that they are exposed frequently to mispronunciations of non-native speakers as well as the fact that many items they know have two similar names which share common structures.

3.2 Materials and methods

3.2.1 Participants

Sixteen 18-20 (range: 556 - 604 days; mean age 579 days) month old infants participated in this study. Half of them were classified as bilingual Spanish-Catalan (4 girls, 4 boys; mean age 585 days) and the other half was considered monolingual Spanish (4 girls, 4 boys; mean age 574 days). An estimation of language exposure was performed based on the number of hours a week the infants were exposed to either language (using an adaptation of the questionnaire developed by Bosch & Sebastián-Gallés, 2001). If the percentage of one language exceeded 70%, the infant was classified as monolingual. Twenty-two additional infants were tested but had to be excluded from the analysis due to fussiness during the experiment (10), being Catalan dominant (4), exposure to a third language (4) or missing parent ratings of the familiarity of stimuli (4).

3.2.2 Stimuli

Six Spanish cognate items were selected to act as experimental stimuli. In our case cognate items refer to stimuli that share etymological origin in Spanish and Catalan. All 6 experimental stimuli were chosen as highly familiar words from the McArthur-Bates Inventory that the majority of infants between 8 and 30 months of age know. All experimental words contained the vowel [e]. In addition, 6 unfamiliar visual items were chosen to fit visual complexity of the experimental stimuli. None of those additional distracter items appeared in the McArthur-Bates Inventory and they were chosen as items that infants of 18 - 20 months of age would not know by name (see Appendix F for images).
A female native speaker of Spanish (bilingual Spanish-Catalan) recorded natural tokens using child-directed speech. All stimuli were recorded in a sound attenuated recording studio at a sampling rate of 44,100Hz and were then edited using Audacity®. Isolated tokens of each word were recorded using the standard pronunciation. Two tokens of each word were selected for use during the exposure phase (mean length 881ms). In addition, two tokens of "ohhh" and the Spanish translation for "look" ("mira") were recorded to serve as auditory stimuli for non-labelled items (mean length 1066 ms). For the test phase two carrier phrases were recorded ("Busca el/la X" translates "Find the X" and "¿Donde está el/la X?" translates "Where is the X?") using each experimental stimulus in the standard and shifted pronunciation (shift [e] to [i]). Mean length of the carrier phrases was 634ms and for the target words 609ms (Mean word length for standard pronunciations was 60 ms and for shifted pronunciations 609ms). None of the shifted items formed a legal word in Spanish or Catalan (the shifted version of "pelota" results in "pilota" and is very close to the Catalan translation which is pronounced / pilɔ́tə /, thus it is not equivalent due to differences in the pronunciation of the vowels [o] and [a]). Following methods from White & Aslin, (2011), we selected for half of the experimental items the sentence frames from standard pronunciations and for the other half sentence frames from shifted pronunciations. With the goal of having both standard and shifted pronunciations occur in the same sentence frame we replaced the target word by its shifted or standard counterpart respectively (i.e. in the sentence "busca el perro" perro was replaced by pirro).

3.2.3 Design and procedure

3.2.3.1 Exposure phase

During an initial exposure phase infants saw 3 displays of 4 objects each. Displays were created using Adobe Flash®. In total the 6 familiar experimental items each appeared in two of the three displays. 3 of those 6 items were labelled with their name in the standard pronunciation while the other half was labelled with tokens of "ohhh" and "mira". Which items were labelled was counterbalanced across two groups between subjects. Each display was shown for 22 seconds. After an initial 2 seconds labelling started and thereafter every 2 seconds a new labelling event occurred (a total of 10 per display). While being labelled, the object loomed on the screen to clarify the relationship between label and object. During each display 2 of the 4 objects were labelled with their Spanish name in the standard pronunciation a total of 4 times, while the other two objects were labelled with "mira" and "ohh" only once. Unlabelled items were included in the displays to be visually familiar, but no information about the pronunciation of these items by the particular speaker was given (Figure 5 illustrates the whole experimental procedure).
Labelled items were shown in a total of 2 displays looming and being labelled a total of 8 times. Unlabelled items appeared in the same number of displays but loomed only twice in total and were accompanied by "ohhh" or "mira". In between displays a dark screen was shown followed by an attention getter (moving colourful circles, accompanied by attractive sounds). Attention-getters would repeat 3 times or until the experimenter judged that the baby was looking.
All participants were exposed to standard pronunciations during the exposure phase. Two lists were created. In list 1 the labelled items were *Chaqueta (jacket), Coche (car)* and *Pelota (ball),* while for list 2 the labelled items were *Pastel (cake), Pera (pear)* and *Vestido (dress).* Accordingly, shifted versions were *Chaquita, Cochi, Pilota, Pastil, Pira* and *Vistido.*

3.2.3.2 Test phase

Right after the exposure displays the test phase began. Static images of one familiar (labelled or unlabelled) experimental item and one of the unfamiliar items (not seen during exposure) were shown. The same unfamiliar object was always paired with the same familiar object for all participants (half of the time the familiar object was on the right and the other half it was on the left).
A total of 24 trials were shown to each infant. 12 trials contained the carrier sentence "Busca el/la X" and 12 contained the sentence "¿Donde está el/la X?". Each of the 6 familiar items appeared in 4 trials: Twice with standard and twice with shifted pronunciation combined with each of the two carrier phrases.
Each trial consisted of a silent baseline phase of 2.5 seconds where the two objects were shown but no auditory stimuli were played. This was done to establish a baseline looking preference for each object pair. After 2.5 seconds the carrier phrase started playing and the display remained on the screen for another 5.5 seconds after sentence onset.
The order of trials was random with the restriction that each carrier sentence could only appear a maximum of three times in a row, the same object could appear no more than twice in a row and accented as well as standard pronunciations could only appear a maximum of 4 of the same type in a row.

3.2.3.3 Procedure

During the experiment the parent sat with the child on her/his lap holding it at the hips. Parents were wearing large dark sunglasses with black tape to avoid interference with the eye-tracker recordings. The Tobii TX 60 eye-tracker screen (51cm x 32cm) was placed at a distance of 50-60cm of the infant's eyes. The display on the screen had a size of 22cm by 16cm. Above the screen was a Sony HDV1080i camera connected to a Macintosh Mac mini computer which was outside of the testing-booth. The experimenter monitored the infant's reaction through the camera. The Tobii eye-tracker was connected to a Macintosh Mac Pro 4.1 computer and controlled by the experimenter from outside the booth. Stimulus presentation was controlled through Tobii studio 2.2.4® software.

Eye data was recorded by the Tobii eye-tracker and software for later analysis.
After an initial calibration using Tobii infant calibration, the exposure phase began. The three displays proceeded automatically with an attention-getter after each display. In total the exposure phase lasted about 1 minutes and 12 seconds assuming 2 seconds between each display.
The test phase began right after the exposure phase with the pseudo random presentation of the 24 test-trials. An attention-getter was placed after every 2-4 trials and was played when the experimenter judged that the baby was not looking (for similar procedure see McMurray & Aslin, 2004). Each trial consisted of a silent presentation of the visual stimuli (pair of one familiar and one unfamiliar object) that lasted 2.5 seconds. After that one of the carrier phrases was played and the display stayed on for another 5.5 seconds after sentence onset; total trial length was 8 seconds. Thus, the test-phase lasted a total of 3 minutes and 40 second assuming no attention-getter was played.
After completing the testing procedure parents were given a questionnaire about the language exposure of the infant, estimating how many hours a week the infant was exposed to the different languages. In addition, another questionnaire assessed parent's judgements about how well the infants knew the experimental items on a scale of 1 (not familiar), 2 (unsure) to 3 (known).

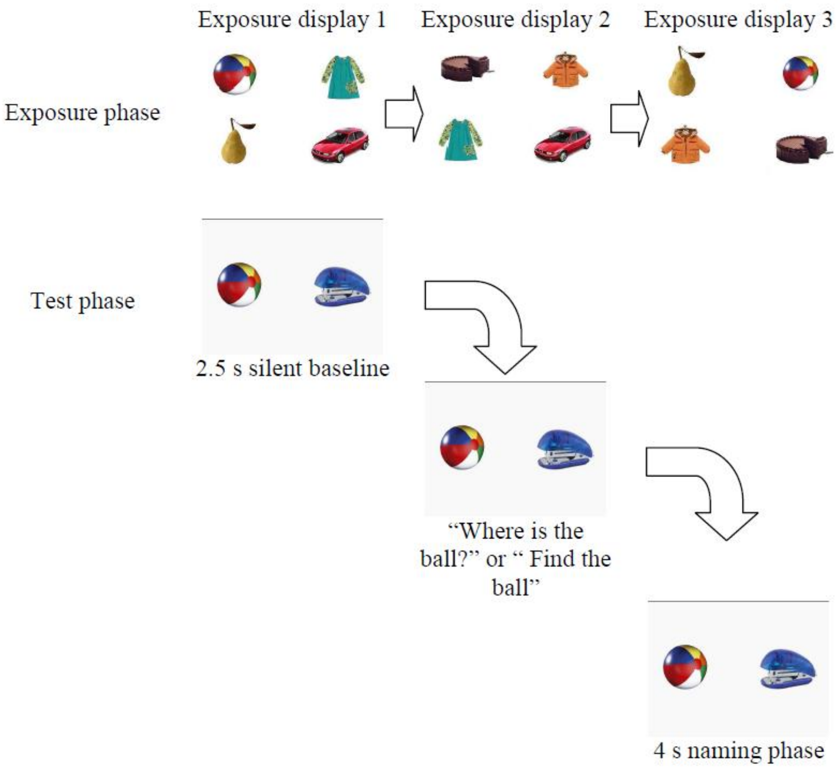

Figure 5 Schematic overview of the procedure during exposure and test phase. Each object was presented twice across the three exposure displays. In the test phase each trial consisted of a silent baseline, the presentation of the carrier sentence and the naming phase (silent).

3.2.4 Eye-movement data analysis

Off-line analyses were performed using Tobii studio 2.2.4® software. Fragments for baseline (onset of visual display until onset of carrier phrase) and naming phase (from 300ms to 4300ms after target word onset) for each trial were extracted and looking times (fixation duration) to both target and distracter stimuli were analysed. The delay of 300ms was used based on previous studies such as White & Aslin, (2011). It has been suggested that this is

approximately the time that is necessary to prepare for and execute saccades (see White & Aslin for a more extensive explanation). Two areas of interest specified in the Tobii software were defined by dividing the screen into left and right. Data were averaged across all trials in each of the four conditions: ***lacc*** (labelled during exposure, shifted pronunciation at test), ***lst*** (labelled during exposure, standard pronunciation at test), ***nlacc*** (not labelled during exposure, shifted pronunciation at test) and ***nlst*** (not labelled during exposure, standard pronunciation at test). Proportion of looks to the target were calculated over the total looking time to target and distracter for both baseline and naming phases of each trial.
Only trials where the infant looked at both objects during the silent baseline phase were included. Furthermore, trials based on target items that parents judged to be unknown or for which they were unsure were excluded from analysis. Infants, who did not have at least one useful trial out of six in each condition or 10 good trials out of 24 in total were excluded from analysis.

3.3 Results

First of all, we compared bilinguals and monolinguals in terms of how many items parents judged to be known. Bilinguals knew an average of 4.75 items out of 6, while monolinguals knew an average of 4.62 items out of 6. No difference between groups was detected ($t<1$). In this analysis we considered only those items that parents judged to be known (3 on our scale).
To ensure that both groups did not differ in their baseline looking preference we looked at the relative percentage of looks to the target during the silent baseline. We compared both groups and found no difference in preference ($t<1$; bilinguals: 47%; monolinguals: 48%).
We calculated relative percentage of looks to the target for both baseline and naming phase and subtracted baseline from naming phase values to obtain the percentage increase in looks to the target after naming. An ANOVA with group (bilingual, monolingual) as between subjects factor and word type (labelled, unlabelled) and test pronunciation (standard, accented) as within subject factors on the increase in relative percentage of looks to the target object was performed. The factor word type (labelled, unlabelled) was marginally significant ($F_{(1,14)}=3.130$; $p=0.099$), while no other main effect reached significance. A marginally significant 3-way

interaction of group by word type by test pronunciation was detected ($F_{(1,14)}$=3.460; p=0.084). In the following we looked at both groups of participants separately and ran an ANOVA including word type (labelled, unlabelled) and test pronunciation (standard, accented) as within subject factors. In neither group did we find any significant effects or interactions. Testing the percentage increase in all four conditions against zero in both groups revealed that bilinguals as well as monolinguals displayed a significant looking time increase only for the ***lacc*** (labelled and accented pronunciation at test) condition (bilinguals: 22% $t_{(14)}$=2.641; p=0.033; monolinguals: 10% $t_{(14)}$=2.354; p=0.051). Comparing bilinguals and monolinguals on each condition did not result in any significant differences for all conditions (Figure 6).

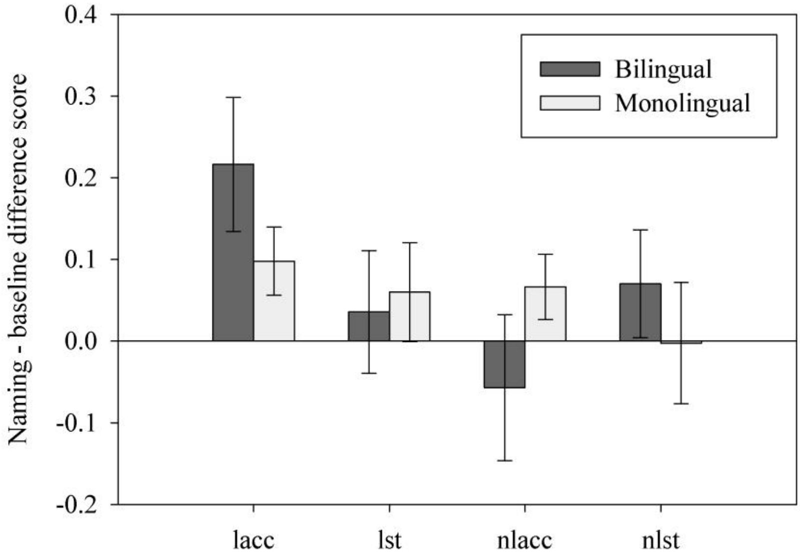

Figure 6 Difference scores for subtraction of percentage of target looks during test phase - percentage of target looks during baseline in all four conditions ± S.E. Only when items were labelled with the standard pronunciation and then produced with an accent at test did infants increase looking to the target after naming. No difference between groups was observed. (lacc=labelled during exposure, with accent at test; lst=labelled during exposure, with standard pronunciation at test; nlacc=not labelled during exposure, with accent at test; nlst=not labelled during exposure, with standard pronunciation at test).

3.4 Discussion

In summary, we found that both groups, bilinguals and monolinguals, significantly increased looking to the target picture only in the condition where the object had been previously named and was then produced with an accent at test. This is a surprising result in the sense that no significant looking increase was detected for objects that were labelled with the exact same name during exposure and at test, which is supposedly the control condition. The fact that only stimuli that parents judged to be known by their infants were used for the analysis suggests that our results are not due to unfamiliarity of the stimuli. Interestingly, the condition in which both groups displayed increased looking to the target after naming was the one that included a change in pronunciation from exposure to test phase (labelled with standard pronunciation during exposure and with accented pronunciation during test). Those results point toward an effect that was not predicted based on previous studies using the same procedure. As is the case for many experiments, infant studies are based on the assumption that the manipulation that is used triggers a certain reaction there is always a possibility that this assumption is not correct. In other words, infants may not necessarily look more to the target item even if they recognise the presented word. In our case, if infants did in fact perform the task as was predicted by earlier studies, we should have found significant looking increases for the ***lst*** condition (standard pronunciation during exposure and test). As this is not the case, results on the other conditions have to be considered with caution. One interpretation that may explain our results is that increased looking in the ***lacc*** condition is based on the mismatch in pronunciation (by the same speaker) between exposure and test. It is possible that we have measured surprise (or confusion) rather than word recognition. As the artificial "accent" used in this study is not a real accent it is possible that due to the inconsistent pronunciation by the same speaker the manipulation was rather confusing and did not induce adaptation or generalisation. In addition, words in the carrier phrases contained the vowel /e/, which was shifted for target words but not the rest of the sentence. Thus, infants could not completely generalise this shift from /e/ to /i/. However, in White and Aslin, (2011) infants also heard the same speaker utter different pronunciations of the same item and still they showed recognition when previously exposed to the shifted pronunciation.

Another tentative explanation for our findings may be that our measure of looking time during the whole 4 second naming phase was not adequate to capture infants´ responses. Those 4 seconds may have been a very long time for infants to look at the same picture and although they might have initially looked at the target object they may have been bored quickly and looked away or to the distracter. The first look after naming would be an interesting measure, but it is not clear at which point in the timeline the infant had recognised the word and was directing its gaze towards the target (or distracter), due to the incremental nature of auditory speech perception. A more adequate way of looking at this data may be to plot individual trial data and see when relatively to the sentence and target word appearance the infants directed their gaze toward one side. Also, it may be more adequate to restrict analysis of the naming phase to the first part as in the total of 4 seconds infants may only "respond" right after naming and then go on exploring the rest of the screen or testing room. In comparison to White & Aslin, (2011) who have used traditional video coding methods, we have used the eye-tracker to record looking times. This may be a crucial difference in terms of precise timing. While during video coding subjective judgement plays a role, the eye-tracker records every gaze point on the areas of interest, which might be an accumulation of very short and not useful looks.

Furthermore, a variety of restrictions made it quite difficult to select the stimuli. First of all, we wanted to use cognates because of the shared structure between languages and the fact that mispronunciations are more likely in those items. Secondly, cognates refer to those objects in a bilingual environment that have another, very similar label and thus, less attention may be paid to the differences. Considering that the vocabulary of infants at that age is relatively small, those restrictions left only a small choice of candidates. Consequently, infants in our study knew overall fewer of the experimental words than in the White & Aslin study, where the average score across items was 3.7 on a scale from 1 (visually unfamiliar and label unfamiliar) to 4 (visually familiar and label highly familiar). Considering that a 3 on this scale meant that the item was visually familiar and the label was familiar, infants in this studies scored very high and most infants new all the words.

Using a similar task as in the present study and eye-tracker data Mulak, et al., (2008) did not find increased target looks when words were produced in an unknown dialect. 19 month old infants saw two familiar pictures and words corresponding to one of the displayed objects were presented in either the native or a foreign accent. Infants only matched the words to the appropriate object when they heard them in their native dialect. Again American English was the native dialect while Jamaican Mesolect was used as a foreign accent. It has to be noted that the Jamaican Mesolect differs strongly from the American English on several dimensions, while the "accent" in White and Aslin, (2011) was a shift in a single vowel. Also it has to be noted that Mulak, et al., (2008) did not use an exposure phase, thus infants encountered this new accent for the first time.

Thus, previous results on accent processing seem to depend strongly on the task and the “accent” manipulations that were used. Tasks using familiarity preference (Best, et al., 2009; Schmale & Seidl, 2009) found that infants as early as 9 months could succeed in the task, while studies using direct word-object matching are more variable. Using a systematic shift in one specific vowel, infants of 19 months were able to match “accented” words to the appropriate object (White & Aslin, 2011), while the use of a strong non native dialect without previous exposure disrupted 19 month olds´ recognition of target items (Mulak, et al., 2008).

The present study presents preliminary findings and needs further testing and analyses to draw conclusions. At the moment, another group of infants is being tested on a version of the experiment using non-cognate items. In addition, a condition using accented pronunciations during the exposure phase is being conducted to control for this factor. At the present time there are still not enough good data in those groups for analyses to be conclusive. Due to our strict criteria on language background and the exclusion of unknown items from the analysis the rejection rate was high and completion of those groups was not possible at this moment.

Thus, more infants are needed on all versions and conditions of the task to be able to draw firm conclusions. In addition, analyses have to be extended to measure the time course of looking to the target item across each individual trial and also across all trials during test, to see the learning effects and possibly capture word recognition more efficiently.

4. NATIVE LANGUAGE EXPERIENCE INFLUENCES RULE LEARNING

4.1 Introduction

Learning a language is a complex process and the nature and scope of the learning mechanisms involved are still a matter of debate. On the one hand, statistical learning has been proposed to be implicated in the extraction of words from continuous speech. On the other hand, rule learning mechanisms help identify non-adjacent dependencies and may ultimately lead to learning of complex grammar and syntax. Both processes have been observed in several studies using artificial languages with controlled transitional probabilities (TPs) and/or imposed rules (Bonatti, et al., 2005; Endress & Bonatti, 2007; Mehler, et al., 2006; Peña, et al., 2002; Saffran, Newport, et al., 1996).
Although those learning mechanisms are universal, meaning that all humans use them, the different languages of the world differ in important properties and, from very early on, characteristics of the native language can significantly shape how speech is parsed. Regarding the phoneme repertoire, newborn infants are initially able to discriminate most phonetic contrasts from many different languages, but by the age of 12 months perceptual tuning has shaped the speech perception system in a way that native contrasts are well perceived, while the ability to perceive contrasts to which infants were not exposed diminishes (Kuhl, 2004; Kuhl, et al., 1992; Polka & Werker, 1994; Werker & Tees, 2005). Even adult listeners tend to assimilate non-native phonemes to their native ones if they sound similar (Best, 1995; Best, McRoberts, & Goodell, 2001; Pallier, Bosch, & Sebastián-Gallés, 1997). Many other examples could be given such as how stress patterns (Cutler & Van Donselaar, 2001; Dupoux, Pallier, Sebastian, & Mehler, 1997; Dupoux, Peperkamp, & Sebastián-Gallés, 2001; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001) or phonotactical rules (Dupoux, et al., 1997; Sebastián-Gallés & Bosch, 2002) of the native language significantly affect how foreign languages are perceived and processed. It has been proposed that, at least for the tuning of the phoneme repertoire, the computation of statistical regularities is a powerful mechanism that can account for native-language tuning processes (Maye, Werker, & Gerken, 2002).

As reviewed below, most studies exploring the effects of properties of the native phonology on statistical learning have focused on the problem of word learning. There is ample evidence indicating that listeners use the knowledge of their native language to constrain the computation of transitional probabilities between syllables. However, to our knowledge, nothing is known about the impact of native phonology on the learning of word-level rules.

4.1.1 Impact of linguistic knowledge on statistical learning

Investigating possible effects of language-specific constrains on word learning, Finn & Hudson Kam, (2008) presented a study showing that adult English speakers did not extract words successfully if they contained consonant clusters at word onset that were not phonotactically legal in English. Similarly, Toro, et al., (2011) showed that statistical learning was disrupted when word forming rules of the native language were violated. Catalan native speakers were exposed to an artificial speech stream containing words that violated Catalan vowel reduction rules (see below for a more extensive description of Catalan phonology). When words from the stream were tested against part-words (spanning word boundaries), participants did not show preference for words. However, when words were tested against new words (not in the stream) or when displayed visually, participants correctly selected words from the stream with a performance above chance. This suggests that although participants were able to segment the stream, they were hindered in selecting the correct item, when it contained a violation of Catalan vowel reduction.
Mixed results were obtained, when testing the influence of word-initial, medial and final stress on word segmentation performance in Spanish, English and French speakers. The interest of testing these three populations was that lexical stress-location differs in an important way. French has final word fixed lexical stress, while English and Spanish show variable stress. However, stress tends to be located on the first syllable in English and on the penultimate in Spanish. As expected, no influences of stress placement on word segmentation were observed for French speakers, for whom stress is not a contrastive feature. However, for both English and Spanish speakers, word segmentation was impaired when attention was

drawn to the medial syllable. Although those two languages differ in stress placement, both groups were affected in the same way. It was concluded that drawing attention away from word boundaries affected English and Spanish speakers, because stress is a contrastive feature in the native language, while French speakers may be deaf to stress cue manipulations (Toro, et al., 2009). Those results provide evidence for an influence of both language specific and more general restrictions.

Those studies have all demonstrated that knowledge of the first language can affect the computation of transitional probabilities to estimate word boundaries under certain circumstances, suggesting a complex interplay of language specific and universal mechanisms. As said, so far all of those studies have considered the effects of linguistic knowledge on word segmentation processes, while we are not aware of data on how word-level rules of the native language may influence rule learning.

4.1.2 Consonants versus vowels (or words versus rules)

Different roles have been attributed to consonants and vowels during information extraction from continuous speech streams. Consonants have been shown to be the unit over which transitional probabilities can be computed to extract word boundaries and thereby words. Vowels on the other hand are at the base of rule extraction, which may translate into syntax and grammar. Bonatti, et al., (2005) showed that when implementing TPs over vowels instead of consonants, participants were not able to extract words. Furthermore, Toro, et al.,(2008) tested the symmetrical hypothesis. When consonants and not vowels carried the structural rule participants were not able to extract and generalise the rule.

4.1.3 The present study

In the present study, we want to explore the effects of knowledge in a first language on rule extraction. Using the same materials in several populations with different language background could help to clarify open issues on the specific influence of a first language on subsequent language learning. Based on the materials used by Toro, Sinnett, & Soto-Faraco, (2011), we created a stream with transitional probabilities between consonants marking word boundaries and imposing an ABA rule over vowels. We selected three groups of participants including Catalan native speakers Spanish native speakers and Hungarian native speakers. We chose those three groups based on the fact that different types of constraints apply to vowels appearing within a word in those languages; specifically, Catalan includes vowel reduction, Spanish has no further rule or restriction on vowels, while Hungarian uses vowel harmony.

4.1.3.1 Vowel rules in Catalan and Hungarian

Most vowels in Catalan are reduced when in an unstressed position. Specifically, the mid front vowels /e/, /ɛ/ and the low central /a/ are reduced to schwa (/ə/) and the mid back vowels /o/ and /ɔ/ are reduced to /u/, while /i/ and /u/ are not reduced. For example, in Spanish the word “mosca” ['m**o**s ka] (fly) and the word “mosquito” [m**o**s 'ki to] (mosquito) are both produced with the same /o/, while in Catalan between the word “mosca” [`m**o**s kə] and the corresponding “mosquit” [m**u**s `kit] the pronunciation of the first vowel changes from /o/ to /u/ due to the shift of stress from the first to the second syllable. In Catalan every word contains one stressed syllable that can hold all vowels but /ə/; vowels in unstressed positions are restricted to /i/, /u/ and /ə/ (Wheeler, 2005).

In Hungarian suffixes are added to word-stems and the vowels contained in those suffixes have to correspond to the vowel in the stem. This means that stems with back vowels (/a/, /á/, /o/, /ó/, /u/ and /ú/) need a suffix containing a back vowel and similarly stems containing front vowels (/e/, /é/,/ö/, /ő/, /ü/, /ű/) have to go with suffixes containing a front vowel to be harmonious (on their own /i/ and /í/ are considered back vowels, while the presence of a front vowel in the stem will turn it into a front stem). Therefore, several versions of each suffix exist and are added according to the vowels in the stem. For example, the dative case can take the back vowel form “nak” or the front vowel form “nek” depending on the stem, so that the stem “bot” will form “botnak” (stick), while the stem ”hit” will form the word “hitnek” (belief).

Important to our hypothesis is the idea that being a native Hungarian speaker rules for vowels are very specific. So, considering the processing of continuous speech, in most situations clear predictions can be made from the vowels in the stem about which vowel will appear in the suffix. In contrast, for native Catalan speakers the rules over vowels are not predictive in the situation of processing continuous speech. Even though the rule restricts vowels in unstressed positions, it is not possible to predict following vowels from those that have appeared previously in the word.

4.1.3.1 Predictions

We have selected three groups of participants with different language backgrounds for the first experiment. Spanish can be considered as the baseline language as no constraints or rules are imposed over vowels. For Catalan speakers restrictions for vowels exist, but no predictions can be made about up-coming vowels. Hungarians however, have learned rules in their native language that make it possible to specifically predict vowels in the end of a word by vowels in the stem.

As stated above, for the present study we have created a continuous speech stream with transitional probabilities between consonants marking word boundaries. In addition, we have imposed an ABA rule over vowels, which means that the first and the last vowel of each CVCVCV non-word are the same, while the vowel in the middle could be one of two options. In other words the last vowel in each non-word could be predicted from the first vowel. We expect that Hungarians will profit from their expectation to find a rule over vowels (which had some parallels with the rule in their native language) and thus, perform best compared to the other groups on the vowel-rule test. Spanish native speakers are expected to perform within the normal range that has been found in previous studies on artificial rule learning and finally, Catalan native speakers may be hindered in learning the ABA rule, as this kind of pattern does not correspond to what was learned in the native language (see Table 3). As for performance on word-segmentation we do not expect any differences between groups as no previous knowledge about constraints for the consonants used in the present study exist in any of the languages.

Table 3 Summary of native language vowel rules and the predicted effect on the extraction of a new vowel rule.

Language	Vowel Rule	Rule	Effect on Vowel rule learning
Catalan	Vowel Reduction	Mid- and low-vowels (/e/, /ɛ/, /o/, /ɔ/, /a/) only in stressed position → only one/word → Rest is reduced	reduced performance
Spanish	No rules	-	-
Hungarian	Vowel Harmony	Stem + suffix → Vowels in suffix correspond to vowels in stem → several versions of each suffix	enhanced performance

4.2 Experiment 1

4.2.1 Materials and methods

4.2.1.1 Participants

In the current study 108 university students participated. Of those, 36 were Catalan native participants from the University Pompeu Fabra in Barcelona, 36 were native Spanish speakers from the Universidad Autónoma of Madrid and 36 were native Hungarian speakers from the Hungarian Academy of Sciences in Budapest. Participants in Barcelona and Madrid received 5 Euro reimbursements, while participants in Budapest received course credit for their participation. Participants from Barcelona were all native speakers of Catalan with early exposure to Spanish. In Madrid participants were monolingual Spanish speakers and in Budapest monolingual Hungarian speakers were selected. A questionnaire evaluating language use was given to all participants. All participants reported some intermediate to low-level knowledge of English learned in school. No hearing or language learning related problems were reported by any participant.

4.2.1.2 Stimuli

We created a continuous artificial speech stream using MBROLA software (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996), using the basic materials of Toro et al., (Toro, Sinnett, et al., 2011). The stream was made up of CVCVCV non-words that were concatenated to form a stream of 10 minutes. Between each word a 25ms pause was inserted to facilitate rule extraction (Peña, et al., 2002). We constructed the non-words by combining three fixed consonant frames (b_d_k, m_l_r, t_p_n) with three vowels (/a/, /i/, /o/) following an ABA rule (see Table 4). All phonemes that were used formed part of the three languages that were tested. Independent native speakers of each language confirmed that the resulting syllables sounded familiar in their own language. The stream was synthesised with the IT4 Italian database, so as to use a language that was not native for any participant. Syllable F0 was set to 240Hz and phoneme duration to 120ms.

We created an artificial speech stream where transitional probabilities (TPs) between the syllables within each word were 0.5, as there were always two possibilities that could follow, whereas between words the TPs were 0.16 as there were six possible syllables that could follow. The between consonant TPs within words were always 1, due to the fixed consonant frames, while between words they were 0.5; Vowels within words had TPs of 0.5 and between words 0.33. It has been shown that using similar paradigms, 10 min of exposure to such TPs in a continuous stream lead to word segmentation and learning of the vowel rule through computation of the given probabilities. For word segmentation, consonants were the important unit to consider, and as TPs were 1 within words and only half as probable between words, statistics could be computed to detect word-boundaries (Toro, Sinnett, et al., 2011).

4.2.1.3 Design and procedure

After listening to 10 minutes of continuous speech stream, participants were presented with an auditory two-alternative forced choice task in the test phase. They heard a pair of stimuli and had to respond which of them was more likely to be part of the language they had heard during familiarisation. One condition tested word segmentation by giving the choice between "words" and "part-words" (part-words are strings contained in the stream, but straddling word-boundaries; see table 2), while the second condition explored rule learning by testing words following the ABA rule against words with an AAB, ABB or AAA structure. For this purpose two new vowels (/i/ and /u/) that had not appeared in the stream were combined with the same three consonant frames that were previously used in the stream, to explore generalisation of the vowel rule to new tokens (see Table 5). Thus, none of the items presented in the second condition had been present in the stream, but half was following the rule and the other half violated the rule. In total, 24 test trials were given; 12 for each condition and the order was pseudo-randomised using three different lists between subjects. The order was random with the restriction that no more than three trials of the same type would be presented successively. Trials consisted of two test items that were separated by a 500ms pause and in half of the trials the correct item was the first and in the other half it was the second item in the test pair. Trials timed out after an eight second delay in case no answer was given. No instructions were given regarding speed of the responses.

At the beginning of the experiment instructions were displayed visually on the screen, telling participants to listen carefully to the extraterrestrial language for a few minutes. After 10 minutes of listening to the stream, instructions for the test phase ("which of the two items is more likely to belong to the language you heard before?") were displayed and participants had the opportunity to ask questions if needed before completing the 24 test trials. Instructions were displayed in the native language for Hungarians and Spanish natives, while Catalan natives were also instructed in Spanish.

For native speakers of Catalan the experiment was controlled using DMDX software and a PC running on Windows. Auditory stimuli were presented over headphones and two keys marked with "1" and "2" (control left and right on the keyboard) were used to respond. Correct and incorrect responses were recorded. Catalan native participants were tested up to four at a time in separate cabins, while for Spanish native speakers all testing conditions were the same except that they were tested up to eight at a time in separate cabins. For Hungarian native speakers testing conditions were the same except for the following. A Macintosh laptop was used to run the experiment and participants were tested one at a time in quite testing facilities.

Table 4 Three consonant frames were combined with three vowels to form the 18 possible words of the "extraterrestrial" language that made up the familiarisation stream.

Consonant frames	Vowels used in the stream		
	A, O	A,I	I,O
B_D_K_	BaDoKa BoDaKo	BiDaKi BaDiKa	BoDiKo BiDoKi
M_L_R_	MaLoRa MoLaRo	MiLaRi MaLiRa	MoLiRo MiLoRi
T_P_N_	TaPoNa ToPaNo	TiPaNi TaPiNa	ToPiNo TiPoNi

Table 5 The two test conditions are displayed below. To test word segmentation words from the stream were presented against part-words from the stream, which straddled word boundaries. Condition 2 tested new tokens following the ABA rule against tokens with an ABB, AAB or AAA rule.

	Test items	
Condition 1 Segmentation using consonants	Words e.g. BaDoKa MoLaRo ToPaNo	Part-words e.g. DoKaMo LoRaTo PaNoBa
Condition 2 Generalisation over vowels	ABA rule e.g. BeDuKe MuLeRu TePuNe	ABB, AAB, AAA e.g. BeDuKu BeDeKu BeDeKe

4.2.2 Results

4.2.2.1 Word segmentation

All groups performed significantly above chance (all $p<0.001$; between 81% and 86.8%) in their preference for words over part-words, Running a one-way ANOVA including all groups no difference was detected ($F<1$). The results showed that in all populations consonants were used successfully to segment the stream into the appropriate words and, importantly, that all populations did not significantly differ in their performance in word segmentation.

4.2.2.2. Vowel rule learning

All groups preferred rule words over non-rule words on the vowel rule condition and thus, successfully generalised the imposed rule to new items. All groups performed significantly above chance with percentages between 56.9 % and 64.8 % (all $p<0.001$).
Due to the fact that we compared the performance of populations across different testing conditions, computer set-ups and participant characteristics (undergraduate or graduate students) we considered performance on word-segmentation as a baseline (for which no cross-language differences were observed). Therefore, between group comparisons were performed on the difference between performance on the consonant minus performance on the vowel conditions.
A one-way ANOVA on the difference score revealed a significant group effect ($F_{(2,107)}=5.120$, $p=0.008$). T-tests indicated lower performance of the Catalan group compared to the Hungarian speakers ($t_{(70)}=3.229$; $p=0.002$). Also, a significant difference was detected when comparing Catalan native speakers and Spanish native speakers ($t_{(70)}=2.222$; $p=0.030$), while no difference for Spanish natives compared to Hungarian natives was found ($t<1$) (Figure 7).

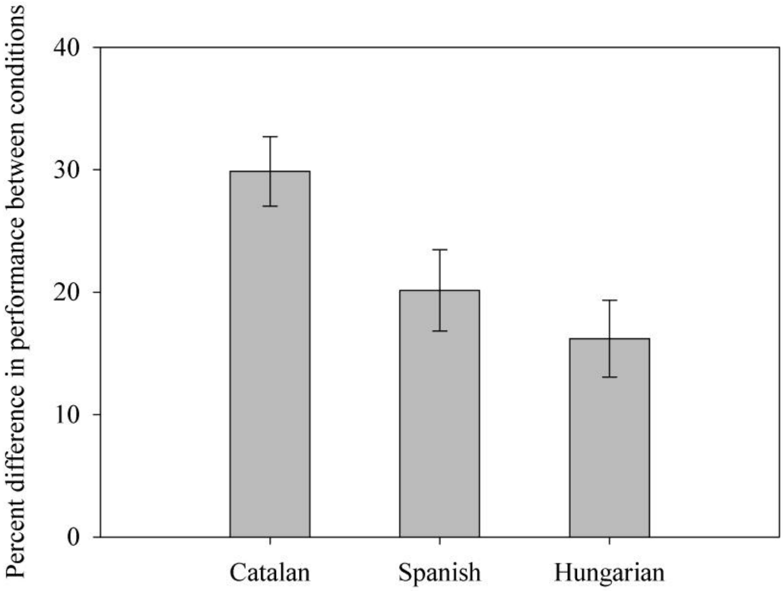

Figure 7 Mean difference score ± S.E. between conditions for all three populations in experiment 1.

4.2.3 Discussion

We have found evidence for an influence of the native language on the extraction of a vowel rule from a continuous speech stream. Confirming our hypothesis, Hungarian natives were best at learning the ABA vowel rule, while Catalan native speakers had the lowest performance. As predicted, Hungarian native speakers have profited from their experience with a similar rule on vowels in their native language, while knowledge of Catalan phonology hindered natives to extract the vowel rule as efficiently.

However, our population of Catalan native speakers was also highly skilled in Spanish as their second language. Thus, it could be possible that in our case the effect found for this group was based on an effect of bilingualism rather than the influence of the native language. To control for the potential influence of knowing a second language in the present results of the Catalan natives, we ran another group of participants, mirror image of the Catalan natives with respect to their knowledge of Catalan and Spanish. The comparison of this group with the native Catalan and native Spanish listeners of experiment 1 should shed light on the issue of a potential influence of knowing Spanish in our group of Catalan natives.

4.3 Experiment 2

4.3.1 Methods

4.3.1.1 Participants

In the present experiment 36 Spanish native speakers from the University Pompeu Fabra in Barcelona participated for a 5 Euro reimbursement. Similarly as for Catalan natives, language use and proficiency was assessed using a questionnaire. All participants had been first exposed to Spanish and had learned Catalan at kindergarten at the latest. Both bilingual groups were from the same environment as in chapter 1. For information on proficiency in Catalan and Spanish of both bilingual groups see Appendix H.

4.3.1.2 Materials and procedure

Materials, procedure and testing conditions were the same as for Catalan native speakers.

4.3.2 Results and discussion

Spanish native bilinguals performed above chance on the word segmentation (83.5%) and the vowel rule generalisation condition (61.1%; for both conditions $p<0.0001$). Using the difference score between conditions as dependent variable as in the previous study, we compared Spanish native bilinguals to both Catalan native (bilinguals) and Spanish native monolinguals. No difference was seen for the comparison with the Spanish native bilinguals ($t_{(70)}=0.497$; $p=0.621$), however, the comparison with the Catalan natives showed a trend towards significance ($t_{(70)}=1.710$; $p=0.092$) (Figure 8).
The lack of differences between both groups of Spanish natives and the trend towards significance in the comparison of both groups

with a high command of both languages indicates that the pattern of results observed in experiment 1 was reliable and that the influence of a second language, even if learned early in life and with a high command has a residual effect in the type of computations we explored here.

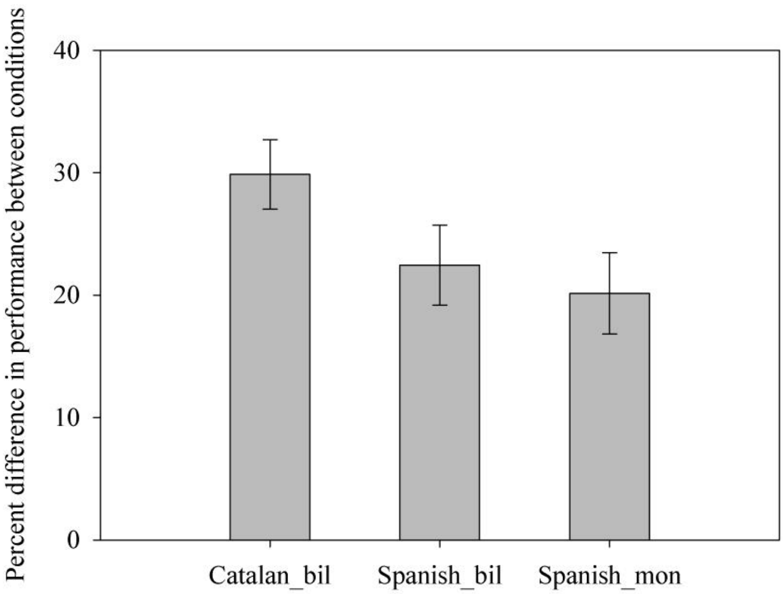

Figure 8 Mean difference score ± S.E. between conditions for Catalan native bilinguals, Spanish native bilinguals and Spanish native monolinguals in experiment 2.

4.4 General discussion

Confirming our hypothesis, Catalan native speakers clearly showed the lowest preference for rule words over non-rule words, while Hungarian monolinguals displayed the highest preference for rule words. Spanish native bilinguals and Spanish native monolinguals performed at an intermediate level between the other two groups.

As explained previously the vowel reduction rule in Catalan is imposing constraints on vowels depending on the stressed syllable.

However, no predictions can be made from the first vowels as to which vowels will appear in the following syllables. In contrast, vowel harmony in Hungarian is based on rules that specifically predict the later vowels from the first vowels occurring in a word.

Our results are in agreement with previous studies that have found strong influences of the first language on the processing of different aspects of phonology even in early and highly skilled bilinguals (for a review see Sebastián-Gallés, 2008). When comparing Spanish native bilinguals to Catalan native bilinguals and Spanish monolinguals, both groups of Spanish natives performed equivalently. A trend towards significant differences between the two bilingual groups indicates that life-long experience with Catalan had a marginal effect on the way Spanish natives processed information concerning vowels. It would have been possible for instance that bilinguals, due to the fact that they have already learned two sets of rules, are less dependent on the influences of the first languages than monolinguals.

Designing materials that equally conformed with the phonological rules in the three languages was a difficult task and some insurmountable problems were encountered. As Spanish has a limited vowel repertoire (five vowels) and all vowels and consonants had to exist in the three languages, it was impossible to create materials that were perfect for all groups. Hence, for both the Catalan and the Hungarian group there were some items that were not following Catalan vowel reduction and Hungarian vowel harmony rules respectively. However, for both groups in half of the 12 trials the correct item was illegal, while in the other half of the trials the incorrect item was illegal[1]. Therefore, it is even more interesting that based on those materials we found evidence of an influence of expectations about rules that were formed by experience with the native language. It may be the case that specific materials adapted to each language could result in larger between

[1] Running by items analyses for each group based on “legal” and “illegal” items showed no effects. However, those analyses have to be considered with caution as they were based on few items and accordingly, do not allow for strong conclusions. A detailed item by item comparison showed that overall (including correct and incorrect items) there were more “illegal” than “legal” items for Hungarians. However, Hungarians still performed best on rule extraction.

group effects (although direct comparisons between groups would be more difficult).

Compared to previous research on the influences of language experience on statistical learning of new words and rules, the present study has provided additional information about the extraction of rules over vowels. It has been previously shown that illegal consonant cluster can disrupt extraction of words from a stream (Finn & Hudson Kam, 2008) and similarly, violating word-forming rules of the native language disrupted selection of appropriate word candidates (Toro, Pons, et al., 2011). Also, language specific stress patterns influenced word-segmentation if stressed positions were manipulated (Toro, et al., 2009). In contrast to all reviewed articles, the present study has not manipulated word-segmentation, but rather the extraction and generalisation of an implemented vowel-rule. Word-segmentation in the present study can be considered a baseline, as all groups segmented the stream into the appropriate words well above chance and differences were found on the generalisation of the vowel rule. Additionally, we used the exact same material to test all populations to contrast specifically expectations about vowel behaviour without any differences in materials. Thus, extending previous research on native language influences on statistical learning the present study provides evidence that although new rules can be learned and generalised, differences in performance based on previous experience with vowels in the native language can be detected. The present results support the notion that although statistical probabilities can be used to extract information about a new language, many factors can guide the listener to pay more or less attention to certain features to restrict the numerous possibilities that tracking statistics might result in (Yang, 2004).

The present results support the idea that basic statistical computations are performed over continuous speech input, while at the same time previous knowledge and expectations guide processing as well. In addition, we extended results on the influence of language experience on statistical learning to the extraction of rules over vowels.

4.4.1Conclusions

In conclusion, we have found evidence that knowledge about rules over vowels in the first language influences they way rules are learned in a new language. Specifically, expectations about the behaviour of certain elements, in our case vowels, seem to have influenced the way rules were extracted and generalised.

5. GENERAL DISCUSSION

5.1 Summary and discussion

The present thesis set out to explore the impacts of the native language environment on auditory speech perception.

In the first series of studies the effects of growing up in a bilingual environment on the processing of indexical variability was studied. Lifelong exposure to mispronunciations and accented utterances in a bilingual environment may have lead to adaptations of the speech perception system to accommodate the increased level of variability. In addition, bilinguals have to manage a larger number of activated lexical entries during auditory speech perception. In the first study we used a recognition memory paradigm with two speakers. Negative effects on error rates when repetitions were uttered by a different speaker than the first presentation were only found for monolinguals. Bilinguals were not affected by the indexical variability inherent in the speaker information. In the second study we increased the level of variability in the environment by introducing more speakers. When participants had to process eight speakers, bilinguals tended to increase their overall performance, while monolinguals got worse. Again, error rates revealed a speaker effect only for monolinguals, while bilinguals were unaffected. The third study used a lexical decision task. No speaker effects were recorded for either group. However, bilinguals performed better and responded faster than monolinguals across all conditions. The difference between groups was especially pronounced on non-words, where bilinguals also showed repetition priming, while monolinguals did not profit from non-word repetition. Taken together, results from the three studies provide evidence that bilinguals were not affected in the same way by indexical variability as monolinguals. Results for the monolingual population replicate previous findings that suggest that fine-grained detail of auditory input is retained and influences priming processes. However, confirming our hypothesis, bilinguals were able to overcome influences from indexical speaker information. Reaction time priming was robust in all three experiments and neither reaction times nor error rates or d-prime were affected by changes in speaker from first presentation to repetition.

To explore the effects of bilingualism on pre-verbal infants´ processing of mispronunciations we ran a study presenting 19 month old infants with a word identification task, including standard pronunciations and shifted pronunciations. Eye-tracking data are still preliminary, but results suggest no difference between bilingual and monolingual infants. Neither group displayed the expected reaction to standard pronunciations and thus, further interpretation of the results were difficult. Both bilingual and monolingual 19 month olds showed increased looking to the target object only for the condition in which items were presented with standard pronunciation during exposure and with the shifted pronunciation at test. This suggests that rather than word identification, we have captured a response that represents surprise or the detection of a mismatch. However, as results are still preliminary and more infants are currently tested, further analysis may shed more light on the original research question of how bilingualism may influence recognition of mispronounced depicted words.

In the third series of studies rule extraction from a continuous speech stream was tested in four groups with different language backgrounds. Rules for vowels in the native language differed between groups and we found an influence of the type of rule in the first language on subsequent learning of a new rule. Hungarian speakers for whom a clear rule for vowels exists that can directly predict vowels in the end of a word by vowels in the beginning of the word learned the imposed ABA rule with the highest performance compared to the other groups. Whereas Catalan native speakers displayed the lowest performance, which was significantly lower than that of Hungarian native speakers. Catalan also has a rule for vowels, but in contrast to Hungarian this rule does not allow for predictions from early vowels in words to later ones. Thus, our results were not merely based on the existence of a rule for vowels but it seems that the similarity between the Hungarian vowel harmony and the chosen ABA rule were at the base of our findings. For native Spanish speakers, who do not have rules over vowels in their native language an intermediate performance was recorded. Furthermore, a second study explored the possible effects of bilingualism using the same materials. Comparing a new group of bilingual Spanish native speakers to both the Catalan bilingual group and the Spanish monolingual group we found that Spanish bilinguals tended to perform in the range of Spanish monolinguals rather than the range of the other bilingual group. This suggests that the measured differences in capacities to extract a rule from an artificial speech stream were mainly influenced by the rules present in the native language rather than being influenced by general effects of bilingualism.

Across three sets of experiments the present work has examined influences of bilingualism and the native language on the processing of auditory speech input. Considering all studies conducted in the present dissertation, we can say that lifelong exposure to a bilingual environment has consequences for general speech perception processes. Adaptations to the increased processing load, such as the accented and mispronounced utterances (Ramon-Casas, et al., 2009; Sebastián-Gallés, et al., 2005; Sebastián-Gallés, et al., 2006; Sebastián-Gallés, et al., 2009) and the non-selective activation of both lexicons (Marian, et al., 2008; Marian & Spivey, 2003; Marian, et al., 2003; Spivey & Marian, 1999), have lead to more flexibility in dealing with irrelevant variability in the speech signal. Although it has to be considered that the effects observed in our studies could be related to general cognitive advantages that have been previously reported for bilinguals (for reviews see Bialystok, 2009; Bialystok, 2010), we propose that the origin may not solely lie in the word selection problem during speech production. Previous results on pre-verbal infants have pointed toward greater flexibility and cognitive advantages of bilinguals, well before the first words were uttered (e.g. Kovács & Mehler, 2009a; Kovács & Mehler, 2009b). However, our results on infants' recognition of mispronounced depicted items are still too preliminary to draw strong conclusions about how being bilingual may affect basic speech perception during language development. Therefore, it is not possible to draw conclusions from the present work about when the adaptations observed in the first set of studies start to shape speech perception during a bilinguals' life.

Furthermore, we have found no effect of bilingualism on rule learning abilities. Given the results of our previous sets of experiments, it might have been possible that bilinguals paid less attention to vowels than to consonants in the artificial language learning experiment, as it has been suggested that vowels carry less relevant information in particular for Spanish-Catalan bilinguals, while common structure between the two languages (corresponding to consonantal frames) may help speed up processing. If this were the case we would have expected bilinguals to perform worse on rule learning irrespective of the first language. Similarly, if generally better cognitive abilities would influence rule learning we would have expected bilinguals to outperform monolinguals. None of this was the case in our study. In contrast, our results suggest that it was mainly knowledge about rules in the first language, which influenced rule learning from continuous speech input.

Taken together, we can say that auditory processing of known languages was affected by bilingual adaptations to their daily input, while processing and rule extraction in a new language was not influenced by bilingualism, but rather by the first language of the participants.

Regarding the question of how and when general speech perception processes are based on abstract or episodic information, we can say that our results support models, which integrate both types of information. The fact that indexical speaker effects only emerged during the slow and effortful recognition memory task, but not when fast lexical decisions were made, gives support to the idea that the time course and level of processing (shallow or deep) are of importance. Furthermore, some hybrid models have tried to explain how both types of information could be integrated.

Goldinger, (2007) has proposed a complimentary-systems approach. In this model speech input passes through primary cortical areas on to the hippocampus where a pre-processed episodic trace is stored in short-term memory (representing episodic information). From the hippocampus the information is than consolidated into long-term memory back in the cortex (representing abstract level) where it can also modify previous knowledge (to account for adaptation etc). Both abstract cortical representations and more or less detailed episodic traces contribute to the decision, depending on time-course and task.

In a slightly different approach, distributed models of auditory speech perception argue that phonology and semantics of the input are processed in parallel vectors, leading to specific patterns of activation in each, that comprise a blend of all possible candidates up to this point in time (Gaskell & Marslen-Wilson, 1997). This necessarily means that up to a words' uniqueness point directly conflicting information is present for lexical candidates as well as for their semantics. At the same time a veridical copy of the input might be available to the system at any point. This enables to have access to all surface detail while at the same time mapping input to more abstract lexical entries.

Both models could explain our data. The fact that both models suggest that at the time of decision a blend of both types of information may be present, supports the idea that bilinguals are better able to manage increased processing loads during speech perception. As we have found that bilinguals were not negatively affected by irrelevant speaker information, while at the same time profiting from non-word repetition, it is plausible that a mechanism that can sort through different types of activations or information was at the base of our findings.

Finally, we can conclude from our results that, indeed, the type of rule inherent in the first language influenced rule learning. While neither an influence of the non-dominant native language, nor a general effect of bilingualism was observed, we can say that the extraction of rules and regularities from unknown speech input was mainly affected by rules that have been acquired very early in life.

This suggests that an effect of bilingualism on this type of rule learning may be detected if simultaneous bilinguals are tested, for whom two sets of rules are present and have to be learned throughout language development. As mentioned earlier, preverbal bilingual, but not monolingual infants at 12 months of age are able to learn two rules simultaneously (Kovács & Mehler, 2009b).

5.2 Conclusions

The present dissertation has provided evidence for an online mechanism that helped adult bilinguals to overcome the increased processing load at the word recognition level, to which they had been exposed throughout their entire life. Furthermore our results support hybrid models of speech perception that integrate both abstract familiarity-based and detailed episodic information during word recognition. Rule learning from an artificial language was mainly affected by linguistic knowledge stemming from the first language. Thus, taking all results together we can conclude that the native language environment shapes the way our speech perception system deals with auditory input. This effect was stronger for native language processing than for foreign language processing.

There are still many open questions that need to be explored in future studies. The proposed mechanism that may be at the base of results from the first series of studies has to be characterised in more detail. One possibility would be to test bilinguals and monolinguals on a similar task, but including responses about the identity of the speaker. This type of experiment could test whether bilinguals could actually make better use of this type of information when it is relevant. If the results of the present work are actually based on a mechanism that sorts through the information present during speech perception to identify relevant activations, bilinguals should also be better at retaining information about the speaker's identity. Furthermore, it would be interesting to explore the neural basis of such a mechanism by using event-related potentials to compare bilinguals and monolinguals.

The presented data on bilingual infants was still preliminary and currently more participants are being tested, as well on a second version of the task using non-cognate items. Different types of further analyses and modifications to the task are possible, such as presenting only the test phase to avoid mismatching pronunciations between exposure and test phase. Furthermore, different talkers could be chosen for standard and shifted pronunciations to better mimic a real accent. We will also take into account more fine-grained details of the recorded results, by looking at individual trial time courses and other measures such as the first look after naming. However, discussing all possible options of modification and analysis for this study exceeds the scope of this discussion.

With respect to influences of native language knowledge on rule learning, it would be very interesting to test more groups with other rules over vowels in the native language to explore more specifically the influence of the exact nature of those rules on subsequent rule learning. It would also be interesting to design materials specifically for each group to observe effects more clearly.

References

Águila, E., Ramon-Casas, M., & Bosch, L. (2007). Patterns of lexical acquisition in bilingual toddlers: Does amount of exposure make a difference? *ISB: 6th international symposium on bilingualism*. Hamburg, Germany.

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science, 14*(2), 395-401.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research* (pp. 167-200). York: Timonium, MD.

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

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words: Research report. *Psychological Science, 20*(5), 539-542.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*(3), 636-644.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism, 12*(1), 3-11.

Bialystok, E. (2010). Bilingualism. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 559-572.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development, 11*(4), 485-508.

Bialystok, E., Craik, F., & Luk, G. (2008a). Cognitive Control and Lexical Access in Younger and Older Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 34*(4), 859-873.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008b). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics, 21*(6), 522-538.

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning Memory and Cognition, 32*(6), 1341-1354.

Bialystok, E., Klein, R., Craik, F. I. M., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290-303.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325-339.

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05) [Computer program].

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science, 16*(6), 451-459.

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics, 39*(4), 514-526.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition, 65*(1), 33-69.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of Early Language Discrimination Abilities in Infants from Bilingual Environments. *Infancy, 2*(1), 29-49.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech, 46*(2-3), 217-243.

Bowers, J. S. (1996). Different perceptual codes support priming for words and pseudowords: Was Morton right all along? *Journal of Experimental Psychology: Learning Memory and Cognition, 22*(6), 1336-1353.

Bowers, J. S., & Kouider, S. (2003). Developing theories on priming with an eye on function. In J. S. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory* (pp. 19-40). Oxford, U.K.: Oxford University Press.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics, 61*(2), 206-219.

Brown, J. S., & Carr, T. H. (1993). Limits on Perceptual Abstraction in Reading: Asymmetric Transfer Between Surface Forms Differing in Typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(6), 1277-1296.

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. E. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *NeuroImage, 17*(3), 1562-1571.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics, 28*(3), 455-474.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science, 12*(5), 815-823.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America, 95*(3), 1570-1580.

Church, B. A., & Schacter, D. L. (1994). Perceptual Specificity of Auditory Priming: Implicit Memory for Voice Intonation and Fundamental Frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 521-533.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116*(6), 3647-3658.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45*(4), 721-736.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic approaches* (pp. 308-324). New York: Oxford University Press.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of

Lexical Access. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*(5), 1283-1296.

Costa, A., Hernandez, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135-149.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59-86.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition, 9*(137-151).

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? *Journal of Memory and Language, 41*(3), 365-397.

Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech, 44*(2), 171-195.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition, 108*(3), 710-718.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*(4), 665-671.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301.

Dorfman, J. (1994). Sublexical Components in Implicit Memory for Novel Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1108-1125.

Duchek, J. M., & Neely, J. H. (1989). A dissociative word-frequency X levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory and Cognition, 17*(2), 148-162.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*(3), 406-421.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *Journal of the Acoustical Society of America, 110*(3 I), 1606-1618.

Dutoit, T., Pagel, V., Pierret, N., Bataille, E., & van der Vrecken, O. (1996). *MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes*.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics, 67*(2), 224-238.

Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica, 135*(2), 182-190.

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition, 105*(2), 247-299.

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development, 78*(5), 1510-1525.

Fennell, C. T., & Werker, J. F. (2003). Early Word Learners' Ability to Access Phonetic Detail in Well-Known Words. *Language and Speech, 46*(2-3), 245-264.

Fernandes, T., Ventura, P., & Kolinsky, R. (2007). Statistical information and coarticulation as cues to word boundaries: A matter of signal quality. *Perception and Psychophysics, 69*(6), 856-864.

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning Memory and Cognition, 32*(5), 1075-1089.

Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition, 108*(2), 477-499.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499-504.

Fiser, J., & Aslin, R. N. (2002). Statistical Learning of Higher-Order Temporal Structure from Visual Shape Sequences. *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(3), 458-467.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276-1293.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*(1), 116-124.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12*(5-6), 613-656.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory.

Journal of Experimental Psychology: Learning Memory and Cognition, 22(5), 1166-1183.

Goldinger, S. D. (1998). Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review*, *105*(2), 251-279.

Goldinger, S. D. (2007). *A complementary-systems approach to abstract and episodic speech perception.* Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, *13*(2), 197-208.

González, J., & McLennan, C. T. (2007). Hemispheric Differences in Indexical Specificity Effects in Spoken Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 410-424.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-81.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1,* 213-219.

Hernandez, A. E., Bates, E. A., & Avila, L. X. (1996). Processing across the language boundary: A cross-modal priming study of Spanish-English bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(4), 846-864.

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism, 13*(3), 315-325.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event Perception and the Word Repetition Effect. *Journal of Experimental Psychology: General, 117*(1), 51-67.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory and Cognition, 12*(6), 568-574.

Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental Science, 12*(1), 131-141.

Jonaitis, E. M., & Saffran, J. R. (2009). Learning harmony: The role of serial statistics. *Cognitive Science, 33*(5), 951-968.

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science, 15*(5), 314-318.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ Detection of the Sound Patterns of Words in Fluent Speech. *Cognitive Psychology, 29*(1), 1-23.

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America, 106*(16), 6556-6560.

Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science, 325*(5940), 611-612.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830-1834.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141-178.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism, 9*(2), 119-135.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831-843.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606-608.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic approaches* (pp. 298-307). New York: Oxford University Press.

Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory and Cognition, 26*(4), 708-715.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1-36.

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition, 74*(3), 347-357.

Marian, V., Blumenfeld, H. K., & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research, 37*(3), 141-170.

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24*(2), 173-193.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language, 86*(1), 70-82.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambrige, MA: MIT Press.

Markman, E. M. (1992). Constraints on word learning: Speculations about the nature, origins and domain specificity. In M. R. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition* (Vol. 25, pp. 59-101). Hillsdale, NJ: Erlbaum.

Martin, C. D., Dering, B., Thomas, E. M., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the 'active' and the 'non-attended' language of early bilinguals. *NeuroImage, 47*(1), 326-333.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science, 32*(3), 543-562.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

McKone, E., & Dennis, C. (2000). Short-term implicit memory: Visual, auditory, and cross-modality priming. *Psychonomic Bulletin and Review, 7*(2), 341-346.

McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(2), 306-321.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of Lexical Form. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*(4), 539-553.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6*(2), 203-229.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex, 42*(6), 846-854.

Merriman, W. E. (1991). The mutual exclusivity bias in children's word learning: A reply to Woodward and Markman. *Developmental Review, 11*(2), 164-191.

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3-4), 1-132.

Mimura, M., Verfaellie, M., & Milberg, W. P. (1997). Repetition priming in an auditory lexical decision task: Effects of lexical status. *Memory and Cognition, 25*(6), 819-825.

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE, 4*(11).

Mulak, K. E., Best, C. T., Irwin, J. R., & Tyler, M. D. (2008). *The effect of dialect on toddler identification of depicted words*. Paper presented at the XVIth International Conference on Infant Studies.

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: A preliminary investigation. *Journal of Child Language, 25*(2), 343-365.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio, 2*, 221-247.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204-238.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1995). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & psychophysics, 57*(7), 989-1001.

Orfanidou, E., Davis, M. H., Ford, M. A., & Marslen-Wilson, W. D. (2011). Perceptual and response components in repetition priming of spoken words and pseudowords. *Quarterly Journal of Experimental Psychology, 64*(1), 96-121.

Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience, 18*(8), 1237-1252.

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

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The Influence of Native-Language Phonology on Lexical Access: Exemplar-Based Versus Abstract Lexical Entries. *Psychological Science, 12*(6), 445-448.

Pallier, C., Dupoux, E., & Jeannin, X. (1997). EXPE: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, and Computers, 29*(3), 322-327.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic Encoding of Voice Attributes and Recognition Memory for Spoken Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 309-328.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604-607.

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology, 19*(3), 395-416.

Philipp, A. M., & Koch, I. (2009). Inhibition in Language Switching: What Is Inhibited When Switching Between Languages in Naming Tasks? *Journal of Experimental Psychology: Learning Memory and Cognition, 35*(5), 1187-1195.

Pisoni, D. B. (1997). Some thoughts on "Normalization" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego: Academic Press.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/o/perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109*(5 I), 2190-2201.

Polka, L., & Werker, J. F. (1994). Developmental Changes in Perception of Nonnative Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 421-435.

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism, 13*(2), 253-262.

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology, 59*(1), 96-121.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26*(4), 392-405.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of Activation. Repetition and Priming Effects in Lexical

Decision and Recognition. *Journal of Experimental Psychology: General*, *114*(4), 435-450.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming test. *Aphasiology, 16*(4-6), 635-645.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*(4), 606-621.

Schacter, D. L., & Church, B. A. (1992). Auditory Priming: Implicit and Explicit Memory for Words and Voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 915-930.

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, *12*(4), 583-601.

Sebastián-Gallés, N. (2008). Cross-Language Speech Perception. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 546-566). Oxford, UK: Blackwell Publishing Ltd.

Sebastián-Gallés, N., & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. [doi:10.1037/0096-1523.28.4.974]. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 974-989.

Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science, 12*(6), 874-887.

Sebastián-Gallés, N., Echeverria, 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., Martí, M. A., Carreiras, M., & Cuetos, F. (2000). *LEXESP: una base de datos informatizada del español*. Barcelona: Universitat de Barcelona.

Sebastián-Gallés, N., Rodríguez-Fornells, A., De Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18*(8), 1277-1291.

Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (2009). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence. *Journal of Cognitive Neuroscience, 21*(12), 2343-2357.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology, 54*(1), 1-32.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*(3), 412-432.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial Activation of an Irrelevant Lexicon. *Psychological Science, 10*(3), 281-284.

Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d-{A figure is presented}/ in monolingual and bilingual acquisition of English. *Cognition, 100*(2), 369-388.

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition, 108*(1), 232-242.

Thiessen, E. D., & Saffran, J. R. (2003). When Cues Collide: Use of Stress and Statistical Cues to Word Boundaries by 7- to 9-Month-Old Infants. *Developmental Psychology, 39*(4), 706-716.

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants: Research article. *Psychological Science, 19*(2), 137-144.

Toro, J. M., Pons, F., Bion, R. A. H., & Sebastián-Gallés, N. (2011). The contribution of language-specific knowledge in the selection of statistically-coherent word candidates. *Journal of Memory and Language, 64*(2), 171-180.

Toro, J. M., Sebastián-Gallés, N., & Mattys, S. L. (2009). The role of perceptual salience during the segmentation of connected speech. *European Journal of Cognitive Psychology, 21*(5), 786-800.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2011). Generalizing linguistic structures under high attention demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 493-501.

Treccani, B., Argyri, E., Sorace, A., & Della Sala, S. (2009). Spatial negative priming in bilingualism. *Psychonomic Bulletin and Review, 16*(2), 320-327.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review, 9*(4), 780-789.

Vila i Moreno, X. F., Vial, S., & Galindo, M. (2004). Language practices in bilingual schools: Some observed, quantitative data from Catalonia. In X. P. Rodriguez-Yañez, A. M. Lorenzo & F. Ramallo (Eds.), *Bilingualism and education: From the family to the school* (pp. 317-329). Munich: Lincoln Europa.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology, 46*(3), 233-251.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science, 14*(2), 372-384.

Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences, 8*(10), 451-456.

Appendix A

Language questionnaire

CUESTIONARIO DE USO DE LAS LENGUAS

Nombre y apellidos..Teléfono......................Código..........
E-mail.. ....
Edad....... Lugar de nacimiento...
Lugar de residencia actual...
si no es donde naciste, desde cuando vives en el lugar actual..................
Lugar de nacimiento del padre...
Abuelos paternos...
Lugar de nacimiento de la madre..
Abuelos maternos...

¿Cuántos años hace que vives en Cataluña? ....
A qué edad comenzaste a escuchar de forma continuada el castellano..............
A qué edad comenzaste a utilizar (hablar) el castellano..............
A qué edad comenzaste a escuchar de forma continuada el catalán..............
A qué edad comenzaste a utilizar (hablar) el catalán..............

a) Indica la lengua (castellano, catalán, ambas u otras) que usualmente utilizas para hablar con:
padre: madre: hermano/as:
novio/a:

b) Si de pequeño hablabas con tus padres o hermanos en alguna otra lengua de la que utilizas actualmente, indica a qué edad se va a producir el cambio:
padre: madre: hermano/as:
novio/a:

c) ¿Qué otras lenguas puedes utilizar (hablar, leer, escribir)?

¿A qué edad iniciaste el aprendizaje formal de éstas lenguas?

Señala la opción que mejor te representa en cada una de las siguientes preguntas:

- Qué nivel de comprensión tienes en estas lenguas:

Francés: perfectamente bien suficiente muy poco
Inglés: perfectamente bien suficiente muy poco
Catalán: perfectamente bien suficiente muy poco
Castellano: perfectamente bien suficiente muy poco

- Qué nivel de lectura tienes en estas lenguas:
Francés: perfectamente bien suficiente muy poco
Inglés: perfectamente bien suficiente muy poco
Catalán: perfectamente bien suficiente muy poco
Castellano: perfectamente bien suficiente muy poco

- Cómo hablas en estas lenguas (fluidez):
Francés: perfectamente bien suficiente muy poco
Inglés: perfectamente bien suficiente muy poco
Catalán: perfectamente bien suficiente muy poco
Castellano: perfectamente bien suficiente muy poco

- Cómo hablas en estas lenguas (corrección de pronunciación):
Francés: perfectamente bien suficiente muy poco
Inglés: perfectamente bien suficiente muy poco
Catalán: perfectamente bien suficiente muy poco
Castellano: perfectamente bien suficiente muy poco

- Cómo escribes en estas lenguas:
Francés: perfectamente bien suficiente muy poco
Inglés: perfectamente bien suficiente muy poco
Catalán: perfectamente bien suficiente muy poco
Castellano: perfectamente bien suficiente muy poco

- ¿En qué lengua te sientes más cómodo/a?
Castellano Catalán En ambas igual

- ¿Si desafortunadamente hubieras sufrido un accidente cerebral que te supusiera la perdida de una lengua, cual de las dos preferirías conservar (sin tener en cuenta criterios prácticos)?
Castellano Catalán

- Si tienes (o tuvieras) un perro o un gato, en qué lengua le hablas (o hablarías)
Castellano Catalán En ambas igual

- ¿Percibes diferencias dialectales?
En castellano *No *Sólo entre algunos dialectos *Sí, claramente

En catalán *No *Sólo entre algunos dialectos *Sí, claramente

- Percibes el contraste entre las vocales “e” abierta (/3/) i “e” cerrada (/e/)?
*Nunca *A veces *Siempre
- Aproximadamente, indica el porcentaje de uso de las lenguas cuando eras pequeño/a:
Castellano........% Catalán........%
¿Y ahora?
Castellano........% Catalán........%
- Has tenido alguna vez o tienes problemas de audición?
En la infancia *No *Si (especifica cual)
Actualmente *No *Si (especifica cual)

- Indica en porcentajes qué mano utilizas para las siguientes actividades

	derecha	izquierda
Escribir		
Dibujar		
Lanzar un objeto		
Cepillarse los dientes		
Usar el cuchillo sin tenedor		
Cortar con las tijeras		
Comer con la cuchara		
Mano más alzada en el palo de la escoba		
Encender una cerilla		
Levantar la tapa de una caja		

Appendix B

Auditory stimuli used in Experiment 1

aceite
acelga
acordeón
águila
ahorro
albaricoque
almendra
anillo
árbol
ardilla
axila
baldosa
basura
bolsillo
bombilla
botella
búho
buñuelo
caja
calabacín
calcetín
caldo
cama
camiseta
cañón
cárcel
carretilla
cebolla
ceja
cepillo
cerebro
cereza
cesta
cielo
ciervo
cigarrillo
ciruela
coche
colchón
comida
conejo
corcho
cosquillas
cruz
cubo
cuchara
cuchillo
cuello
cuerda
cuervo
dedal
demonio
desierto
diciembre
dieciséis
diente
disfraz
duende
enchufe
enfermo
escarabajo
escoba
espalda
espárrago
espejo
falda
fresa
fuente
gafas
globo
grifo
guisante
hielo
hoja
hombre
hormiga
iglesia
índice
jarrón
judía
julio
ladrillo
lechuga
león
leopardo
libro
limón
lluvia
lobo
lunes
maestro
mano
manzana
mariposa
martillo
mechero
médico
melocotón
mesa
mujer
muñeco
naranja
nieve
ojo
órgano
oveja
paloma
pañuelo
pato
pavo
peonza
perejil
perro
pez
pimiento
pincel
puente
puerta
puerto
queso
raíz
rana
ratón
rebanada
red
rojo
ruido
sábana
sombrero
tenedor
tierra
toalla
tobillo
tomate
trenza
uva
vaso
vela
ventana
vírgen
yema
zanahoria
zapato
zorro

Appendix C

Auditory stimuli used in Experiment 2

aceite	cigarrillo	grifo	paloma
acelga	ciruela	guisante	pañuelo
acordeón	coche	hielo	pato
águila	colchón	hoja	peonza
ahorro	comida	hombre	perejil
albaricoque	conejo	hormiga	perro
almendra	cosquillas	índice	pez
anillo	cruz	jarrón	pimiento
ardilla	cubo	judía	pincel
axila	cuchara	julio	puente
baldosa	cuchillo	ladrillo	puerto
basura	cuello	lechuga	raíz
bolsillo	cuerda	león	rana
bombilla	cuervo	leopardo	ratón
botella	dedal	libro	rojo
búho	demonio	lluvia	ruido
caja	desierto	lobo	sábana
calabacín	diciembre	lunes	sombrero
calcetín	diente	mano	tenedor
caldo	disfraz	manzana	tierra
cama	duende	mariposa	toalla
camiseta	enchufe	martillo	tobillo
cañón	enfermo	mechero	tomate
cárcel	escarabajo	médico	uva
carretilla	escoba	melocotón	vaso
cebolla	espárrago	mujer	vela
ceja	espejo	muñeco	ventana
cepillo	falda	naranja	virgen
cerebro	fresa	nieve	yema
cereza	fuente	ojo	zanahoria
cesta	gafas	órgano	zapato
cielo	globo	oveja	zorro

Appendix D
Auditory word stimuli used in Experiment 3

acelga	ciervo	iglesia	pavo
acordeón	ciruela	índice	perejil
ahorro	coche	jarrón	perro
albaricoque	colchón	judía	pez
almendra	comida	julio	pimiento
anillo	cruz	ladrillo	pincel
ardilla	cuchara	lechuga	puerta
baldosa	cuchillo	león	puerto
basura	cuello	leopardo	rana
bolsillo	desierto	libro	ratón
bombilla	diente	limón	ruido
búho	disfraz	lobo	sábana
caja	duende	maestro	sombrero
calabacín	enfermo	mano	tierra
calcetín	escarabajo	manzana	tobillo
caldo	escoba	mariposa	trenza
cama	espárrago	médico	uva
camiseta	espejo	melocotón	vaso
cañón	fresa	naranja	virgen
ceja	gafas	nieve	yema
celdo	globo	ojo	zanahoria
cepillo	guisante	oveja	zorro
cerebro	hielo	paloma	
cereza	hoja	pañuelo	
cielo	hombre	pato	

Auditory non-word stimuli used in Experiment 3

abolado	corsata	linterda	químita
abrito	cosona	llada	richa
amido	desío	mansel	rinoceponte
amposa	dialante	mañeta	sábaco
ancra	dicujo	medocina	sadero
arquidecto	diplora	menoje	salcón
aruna	discofeca	mentica	saliga
asionto	embatido	merentar	sancía
aspana	espatuto	milapro	selma
ballera	estalera	mochina	semina
banduja	fido	monca	silada
bite	fogo	moneca	siñón
buidre	fruca	múscula	sobrida
cafor	fuerpa	naruz	tepado
carnizuría	gabardila	nefro	terraba
carpueta	ganator	nerpio	teza
carsera	grenja	olo	tiberón
célila	guaste	omblago	tijetas
cetar	guitaca	otor	tojo
chiga	hercano	pantasa	tortifa
chimenía	hipitótamo	pelgue	trombán
cirno	hosbital	prequio	ugiptio
code	inverco	primasera	
colepio	lágrina	prograna	
cordere	leta	psicálogo	

Appendix E

Language proficiency of participants in chapter 1. Language skills were assessed after the experiment using a questionnaire. Language scores represent proficiency levels on a 4-point scale (1.0= very bad and 4.0=native level).

		Experiment 1		Experiment 2		Experiment 3	
		Bil	Mon	Bil	Mon	Bil	Mon
S	Comprehension	4.0	4.0	4.0	4.0	3.9	3.7
	Reading	4.0	4.0	4.0	4.0	3	3.6
	Speaking	4.0	4.0	4.0	4.0	3.9	3.7
	Pronunciation	4.0	4.0	4.0	4.0	3.9	-
	Writing	4.0	4.0	4.0	4.0	3.9	3.7
C	Comprehension	4.0	-	4.0	-	3.9	-
	Reading	4.0	-	4.0	-	3.9	-
	Speaking	3.9	-	4.0	-	3.7	-
	Pronunciation	3.9	-	4.0	-	3.6	-
	Writing	3.9	-	4.0	-	3.7	-
E	Comprehension	2.9	2.7	3.2 *	2.9	2.7	2.3
	Reading	3.0	2.8	3.4 *	2.7	2.8	2.2
	Speaking	2.6	2.4	2.9 *	2.5	2.3	1.8
	Pronunciation	2.7	2.4	2.9 *	*	2.3	*
	Writing	2.8	2.5	3.0 *	2.7	2.5	2.0

Bil = Bilingual, Mon= Monolingual; * Missing data from 15 participants; S=Spanish, C=Catalan, E=English

Appendix F

Familiar items	Unfamiliar items
	
	
	
	
	
	

Appendix G

Test pairs for both word segmentation and rule learning tests in Experiment 5 & 6

Word segmentation		Rule learning	
Words	Part-words	Rule words (ABA)	Non-rule words
badika	kotapo	tupenu	beduku
panoma	milari	bedeku	melure
tiponi	ribadi	melere	tepune
tipani	lariba	muleru	tepene
nibadi	malora	tepunu	budeku
bodiko	kitopi	muluru	beduke
topino	nobido	mulere	tepune
lirobi	badoka	tupunu	melure
bidaki	ponimo	tepenu	budeku
dikati	tapona	beduke	meleru
moliro	naboda	tupune	mulero
molaro	raboda	tupenu	buduke

Appendix H

Language proficiency of Catalan native and Spanish native bilinguals in chapter 3. Language skills were assessed after the experiment using a questionnaire. Language scores represent proficiency levels on a 5-point scale (1.0= very bad and 5.0=native level).

		Catalan bilinguals *	Spanish bilinguals #
Spanish	Comprehension	4.8	4.8
	Reading	4.8	4.9
	Speaking	4.6	4.8
	Writing	4.7	4.9
Catalan	Comprehension	5	4.8
	Reading	5	4.6
	Speaking	5	4.3
	Writing	5	4.3

* Missing data from 11 participants; # Missing data from 3 participants