—Grau
en Llengües
Aplicades

Treball de Fi de Grau

Categorical speech perception in bilinguals:

A review on previous research and an experimental proposal

Michelle Throssell

Tutor: Mario Bisiada

Seminari 101: Descripció i Comparació de Llengües

Curs 2019-2020

upf. Universitat Pompeu Fabra Barcelona

Facultat de Traducció i Ciències del Llenguatge

ACKNOWLEDGEMENTS

First of all, I would like to thank my tutor, Mario Bisiada, for accompanying and advising me throughout this process and to Sílvia Perpiñán for her classes on *Language Acquisition* and resolving my doubts regardlessly.

I would like to express my gratitude to my *Languages and the Mind* and *Cognition and Language* professors at Pompeu Fabra University, Helena Roquet, Aurora Bel and Sara Cañas for awakening my curiosity about language acquisition and the cognitive processes involved in this capacity.

I would also like to thank João Veríssimo for his fascinating lessons during my exchange year at the University of Potsdam, in one of which he presented the research on which my proposal is based and by which is inspired and for giving me feedback on my design.

My study would not have been possible without all investigators that have given me free access to their research, such as Núria Sebastián-Gallés, Laura Bosch and Janet Werker, among others.

Lastly, I wish to thank my family and friends for listening to and bearing with my trivial complaints about writing this project, the pandemic and all its complications.

ABSTRACT

The study of the acquisition of categorical perception of speech has a complex trajectory. A large number of factors must be taken into account, and those are often impossible to control. Earlier research used to focus primarily on the first steps of language acquisition in monolinguals, while in recent years the study of the same abilities in bilinguals has been preferred. The methodology for assessing the discrimination capacity between contrastive sound pairs has also transformed drastically. Regarding the determining factors for the phonetic contrast differentiation task, new possibilities are added in each study, especially in the bilingual spectrum. In bilingual research, not only the frequency of each sound in everyday language has to be taken into consideration, but also the proportionality of each language in everyday life. In the present work, I review less studied factors proposed by different researchers as possible variables influencing bilingual discrimination, such as the ratio of cognates shared by many language pairs and the nature of the contrastive sounds to be evaluated. Finally, I offer an experimental proposal to study the development of this phenomenon in bilinguals who differ in the number of cognates shared by their two native languages.

Keywords: speech perception, language acquisition, bilingualism, phonological discrimination

RESUM

L'estudi de l'adquisició de la percepció categòrica del llenguatge té una trajectòria complexa. S'han de tenir en compte una gran quantitat de factors que moltes vegades són impossibles de controlar. Als inicis d'aquest àmbit, la recerca se centrava especialment en els primers passos de l'adquisició del llenguatge en monolingües, mentre que en els últims anys s'ha preferit l'estudi de les mateixes habilitats en bilingües. La metodologia per avaluar la capacitat de discriminació de sons contrastius també s'ha transformat dràsticament. Respecte als factors determinants per a la tasca de diferenciar contrastos fonètics, s'afegeixen a cada estudi noves possibilitats sobretot en l'espectre bilingüe. En la recerca bilingüe, no sols s'ha de tenir en compte la freqüència de cada so en el llenguatge quotidià, sinó que també la proporcionalitat que suposa cadascuna de les llengües al dia a dia. En el present treball, es revisen també factors menys estudiats però proposats per diferents investigadors com a possibles variables influents en la discriminació bilingüe, com la ràtio de cognats que comparteixen molts parells de llengües i la naturalesa dels sons contrastius a avaluar. Finalment, s'ofereix una proposta experimental per estudiar el desenvolupament d'aquest fenomen en bilingües que difereixen en la quantitat de cognats compartits per les seves dues llengües nadiues.

Paraules clau: percepció de la parla, adquisició del llenguatge, bilingüisme, discriminació fonològica

RESUMEN

El estudio de la adquisición de la percepción categórica del lenguaje tiene una trayectoria compleja. Se deben de tener en cuenta una gran cantidad de factores que muchas veces son imposibles de controlar. En los inicios de este ámbito, la investigación se centraba especialmente en los primeros pasos de la adquisición del lenguaje en monolingües, mientras que en los últimos años se ha preferido el estudio de las mismas habilidades en bilingües. La metodología para evaluar la capacidad de discriminación de sonidos contrastivos también se ha transformado drásticamente. En cuanto a los factores determinantes para la tarea de diferenciar contrastes fonéticos, se suman en cada estudio nuevas posibilidades sobre todo en el espectro bilingüe. En la investigación bilingüe, no solo se tiene que tener en cuenta la frecuencia de cada sonido en el lenguaje cotidiano, sino que también la proporcionalidad que supone cada una de las lenguas en el día a día. En el presente trabajo, se revisan también factores menos estudiados pero propuestos por diferentes investigadores como posibles variables influyentes en la discriminación bilingüe, como el ratio de cognados que comparten muchos pares de lenguas y la naturaleza de los sonidos contrastivos a evaluar. Finalmente, se ofrece una propuesta experimental para estudiar el desarrollo de este fenómeno en bilingües que difieren en la cantidad de cognados compartidos por sus dos lenguas nativas.

Palabras clave: percepción del habla, adquisición del lenguaje, bilingüismo, discriminación fonológica

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Motivation and interest

I have always been fascinated about languages: their complexity, their history, their power and their differences and similarities. There is nothing related to language that is not worth noting, or at least that does not draw my attention. Nevertheless, during my studies in Applied Languages, I have discovered what is for me the most intriguing field of study within the scope of linguistics: psycholinguistics.

The cognitive processes behind language acquisition and development are captivating. Newborns can recognize their own language only from the exposure received during their months in their mother's womb. Children so young that still do not know how to tie their shoes are able to comprehend and produce language and sometimes even more than just one. I wonder how is it possible that this scientific field is relatively new and that such phenomenons were taken for granted till not that long ago. However, I am grateful for those who have been researching these issues for decades, and I am willing to become one of them in the near future.

The specific matter I would like to analyze is the development of phonological categorization in bilinguals. Although this subject has been previously studied with monolinguals, the literature about simultaneous bilinguals is scarcer. Furthermore, the findings on speech perception in bilinguals are not compelling. The most remarkable conditions that may be leading to these mixed results are a) the adopted method (preference-familiarization vs. Anticipatory Eye Movements), b) the number of cognates shared within the two languages and c) the nature of the phonemes tested (vowel discrimination vs. consonant discrimination).

My intention with this project is to offer a research proposal that overcomes these inconveniences. I am especially enthusiastic about this study design, because it requires a familiar linguistic sample, while the vast majority of studies are usually based on the English language. Even though it entails a large population and therefore offers many opportunities to carry out experimental tasks, exclusively relying on English is a great mistake that could lead to just partial findings. I reckon that linguistic diversity is crucial to develop a broader perspective and conduct studies that lead to more significant results.

1.2. Introduction to the study and research questions

As listeners, we are trained to filter the acoustic variation we receive and to perceive and identify clearly separated categories. This phenomenon is known as categorical perception. The acoustic continuum is divided according to the distinctive sounds of our language. The acquisition of categorical perception is one of the first milestones in language acquisition. Toddlers between the ages of 6 and 8 months begin to have difficulty discriminating minimal pairs that do not constitute phonological contrasts in the mother tongue, while they distinguish relatively well those that are. Thus, the phonemic system of a language will strongly condition both the production and the perception of linguistic sounds. When hearing a phonic sequence, listeners are predisposed to look for acoustic cues that help us perceive the distinctive units and ignore the rest of phonetic variations. This development can be studied by measuring until what age infants can discern between non-contrasting sounds in their language since the loss of this capacity evidences the formation of phonological categories.

A variety of studies have tried to answer the questions that arise from these statements: what are the consequences of being exposed to more than one language from the first months of life if they do not share the same phonemic system? Will it be enough for a phonetic contrast to be present in one of the languages? Will there be differences from monolingual peers in this development? However, the results have not been conclusive. While some studies show a different pattern in the ability to discriminate language-specific contrasts among bilinguals, other studies do not replicate such distinction. Instead, they show that this capacity remains the same throughout the developmental process.

In the present work, the existing research on both monolingual and bilingual categorical perception acquisition will be synthesized. Secondly, the issues that seem to have driven this field to the mentioned above discrepancies, such as the methodology, the ratio of cognates and the nature of the input, will be reviewed. Subsequently, the linguistic environments of the Basque Country and Catalonia will be described, as Basque and Catalan are the selected languages for the following proposal. Lastly, a detailed experimental design and the plausible experimental results will be presented and discussed. This study would focus on a novel question within the field: Are there differences in language acquisition development between different groups of bilinguals?

2. BACKGROUND

2.1. Synthesis of existing studies on speech perception in monolinguals

This section synthesizes some of the studies regarding the perception of speech in monolinguals in the following areas: language discrimination, the establishment of the phonetic inventory and the audiovisual perception of speech. As currently the amount of literature in this area is relatively extensive, the main findings will be shared through the earliest pieces of work, while also mentioning more modern research.

2.1.1. Language discrimination

Identifying one's language is the very first step toward language development. According to a series of studies by Mehler et al. (1988), infants can discriminate between languages based on their rhythm. This ability was tested by recording the sucking responses (high-amplitude sucking procedure[1]) of 4-day-old French and 2-month-old American babies while they were listening to sequences in different languages. The results indicated that the infants were able to discern between their native language and another language (French from Russian for the French newborns and English from Italian for the American toddlers), but not between two foreign languages (French from Russian for the American toddlers and English from Italian for the French newborns). The first main conclusion was that in order to be able to distinguish two languages, there must be some familiarization with one of them. Over the course of a series of tests with filtered versions of the sequences, it was later established that infants relied on prosodic information in order to discriminate their native language from another, provided that the other had a different rhythm.

A similar investigation (Nazzi, Bertoncini & Mehler, 1998) tested this capacity with French newborns within the first 5 days of life. Throughout the first and second experiments, it was established that infants discriminated English from Japanese, but not English from Dutch. The main prosodic difference between these languages is that both English and Dutch have a stress-timed rhythm, while the Japanese rhythm is mora-timed. In a third experiment, different combinations of sequences in English, Dutch, Spanish and Italian were used. Newborns were only able to discriminate between languages when rhythm types were

[1]Infants are given a pacifier that is attached to a computer which measures changes in air pressure. Newborns learn that high-amplitude sucking results in auditory stimuli. Thus, they suck more intensively when they are attracted to the stimuli presented.

contrasted, i.e. English or Dutch (stress-timed) against Spanish or Italian (syllable-timed). These results are consistent with those of Mehler et al. (1988), providing more extensive data on the capacity of infants to discriminate languages by their rhythm.

Further research, carried out by Ramus, Nespor and Mehler (1999), is consistent with the reviewed above studies but note that those do not prove that infants discern languages relying on rhythm alone as other types of prosodic information are preserved in the auditory stimuli. In addition, more recent research by Chong, Vicenik and Sundara (2018) provided evidence that babies are able to distinguish languages from the same rhythmic class based on other prosodic cues, such as intonation. On the other hand, according to the investigation conducted by Nazzi, Jusczyk and Johnson (2000), infants were able to discriminate languages from the same rhythmic category when the contrast included the familiar language but not when they were faced with different foreign languages. Thus, retrieving the first conclusion from Mehler et al. (1988) about the role of familiarity.

2.1.2. Establishment of the phonetic inventory

A pioneering investigation regarding infant speech perception had already shown that newborns are able to discriminate between phonetic contrasts (Eimas et al., 1971). Earlier research had established the ability to discern between the voiced and voiceless forms of the stop consonants /b-p/, /d-t/ and /g-k/ in adults. In their study, this same capacity was tested with babies from 1 to 4 months old. Phonological discrimination was assessed through the high-amplitude sucking procedure while listening to a phonetic continuum containing the mentioned above contrasts. Once the participants were habituated with a particular sound, the sound varied (for testing condition) or not (for control condition). When the sucking rate increased and longer times were needed to recover the habituation sucking rate, the ability to discern phonemes was evidenced. The results indicated that toddlers perceive phonetic contrasts regardless of having had little exposure to language, suggesting that the categorical perception of speech from an extraordinarily early age.

Along these lines, Streeter (1976) tested 2-month-old Kikuyu[2] newborns on their capacity to discriminate labial stop consonants based on voicing. Phonemes such as /b/ and /p/ differ on their onset voicing time. In /p/ voicing (vibration of the vocal cords) follows release

[2]Bantu ethnic group inhabiting Central Kenya. Kikuyu also refers to their native language, which is a member of the Bantu subgroup of the Niger-Congo language family.

burst, while in /b/ voicing and release occur almost simultaneously. In this study, reactions to three different onset voicing times were analyzed: pre-voiced, voiced and voiceless. As in Eimas et al. (1971), phonological discrimination was assessed by measuring the sucking reactions among the infants while listening to the varying stimuli. It was found that babies were able to detect differences in voicing onset times even in the absence of linguistic exposure in which these categories are relevant. Different from Eimas et al. (1971), the results indicated discrimination for pre-voiced vs. voiced conditions, in addition to the voiced vs. voiceless distinction.

Werker et al. (1981) studied the potential of this ability during the first months of life and the possibility of a rapid decline. The capacity to discriminate two Hindi speech contrasts among adult English speakers, adult Hindi speakers and 7-month-old infants was compared. For this purpose, the participants were assessed through the head-turn preference procedure[3], while they were presented with the English speech contrast /b-d/ and two contrasts from Hindi (not present in English). The first Hindi contrast was the distinction between the voiceless unaspirated retroflex /T/ and the dental stop /t/, while the second contrast was the distinction between the voiceless aspirated dental stop /t^h/ and the voiced aspirated dental stop /d^h/. According to the results, all infants were able to discern not only the English contrast but between the non-contrastive sounds, while only Hindi adults could distinguish these sounds. These findings evidenced that babies have broader perceptive abilities that develop into a language-specific capacity.

Along these lines, Werker and Tees (1984) attempted to determine the generalizability of this process and delineate its time course. In their study, English infants ranging from 6 to 8 months, Thompson[4]-speaking adults and English-speaking adults (that had not had exposure to a second language containing the contrast being studied) were tested on their ability to discriminate different phonological contrasts from both English and Thompson through the same procedure. The sound pairs were stop consonants diverging in their place of articulation: /ba/-/da/ (bilabial vs. alveolar) as the English contrast and /ki/-/qi/ (velar vs. uvular) as the Thompson contrast (not present in English). According to the criterion for discriminability established by the researchers, all Thompson-speaking adults, 8 out of 10

[3]Interest is measured by recording the amount of time that babies attend to the stimuli presented. By comparing infants' preference among two or more stimuli types, researchers can determine whether babies are sensitive to the relevant properties that differentiate the stimuli.
[4]The Thompson language is an Interior Salish (Native Indian) language spoken in south-central British Columbia.

infants but only 3 out of 10 English-speaking adults were able to discern the Thompson contrast successfully, replicating the previous findings. A second experiment, compared infants aged from 8 to 10 months and 10 to 12 months to the earlier data collected on toddlers aged from 6 to 8 months under identical testing conditions. Furthermore, a third longitudinal study tested the same 6 babies at three ages: 6-8 months, 8-10 months and 10-12 months old. The achieved results lead to the conclusion that the decline in the capacity to discriminate non-contrastive sounds occurs progressively during the first year of life.

Latter investigations have provided further evidence on the ability of infants to differentiate phonemes even when those contrasts do not belong to the phonetic inventory of their language and that this capacity critically deteriorates for non-native contrasts, while it significantly improves for native contrasts (Kuhl et al., 2006; Narayan, Werker & Beddor, 2010). Furthermore, the studies conducted by Kuhl et al. (1992) and Polka and Werker (1994) showed that this pattern in acquisition also occurs for vowel contrasts. More recent research has studied infant phonetic discrimination by using EGG[5] and MEG[6] measures, allowing the investigation of the relationships between brain structures and their functions and providing more substantial evidence for the previous findings (reviewed by Gervain & Mehler, 2010; Kuhl, 2010).

2.1.3. Audiovisual perception of speech

Another relevant factor that should not be neglected regarding speech perception is that babies are sensitive to both auditory and visual stimulation during language development (Kuhl & Meltzoff, 1982). In this study, infants ranging from 4.5 to 5 months were presented with a silent film of two speakers, one of them articulating /a/ and the other articulating /i/, while a track of one of these sounds was displayed. By scoring the infants' visual fixations to the faces, it was found that participants tended to perform longer looking times for the face that matched the audio. This phenomenon is not exclusive to early infancy, as it has been found that it is easier to understand a masked voice when the visual information of the face is simultaneously visible (Sumby & Pollack, 1954).

Further evidence regarding this matter is the so-called McGurk effect (McGurk & MacDonald, 1976), a perceptual illusion that occurs due to the interaction between auditory

[5]Non-invasive technique that records the brain's electrical activity through the use of electrodes.
[6]Non-invasive technique that records functional brain activity by capturing magnetic fields.

and visual stimuli. In their experiment, participants were presented with a film of a speaker producing the syllable /ga/, whose audio was dubbed into a /ba/ syllable. Surprisingly, most participants reported that they heard the syllable /da/. This phenomenon occurs when the auditory component of one sound is simultaneously presented with the visual component of another sound, resulting in the perception of a third sound. Rosenblum, Schmuckler and Johnson (1997) tested the McGurk effect with 5-month-old English infants and showed that the illusion can be triggered from early infancy.

Lewkowicz and Hansen-Tift (2012) found that infants between 4 and 8 months old switch from looking at the speaker's eyes to their mouth just before their first babbles. In their investigation, the gaze of babies ranging from 4 to 12 months was traced while they watched and listened to a female reciting a monologue either in their native language (English) or a non-native language (Spanish). Between the ages of 4 and 8 months old, the newborns' attention shifted from the eyes to the mouth of the speaker regardless of the language spoken. At 12 months old, their gaze switched back to the eyes of the talker, but only for the native language. The researchers stated that the cause of this difference is the narrowing of speech perception that occurs during the first year of life. Just like in the very beginning, 12-month-olds still draw their attention to the speaker's mouth, searching for visual cues that help them decipher what now has become a foreign language.

2.2. Synthesis of existing studies on speech perception in simultaneous bilinguals

This section synthesizes some of the studies regarding the perception of speech in bilinguals in the same areas as the previous one about monolinguals: language discrimination, the establishment of the phonetic inventory and the audiovisual perception of speech. It closely follows the pioneering work by Bosch and Sebastián-Gallés (1997, 2001, 2003a, 2005, 2009). In addition to the study conducted by Albareda-Castellot, Pons and Sebastián-Gallés (2011), on which my proposal is based and by which it is inspired.

2.2.1. Language discrimination

The earlier suggestions that newborns can discriminate between a familiar and a foreign language when those have a different rhythm raised some questions about this ability in bilinguals. Would they be able to discern between two familiar languages within the same rhythm category? A pioneering study regarding this matter is a set of experiments with 4-month-old Catalan and Spanish monolingual and bilingual infants (Bosch & Sebastián-

Gallés, 1997). After testing the effectivity of a visual orientation procedure[7] with a pair of phonologically dissimilar languages (Catalan or Spanish vs. English), it was found that bilinguals use a different strategy than monolinguals to discriminate between their native languages and a foreign language (English or Italian). While monolinguals oriented faster to the familiar language, bilinguals oriented more rapidly to the unfamiliar one. In addition to this reversed pattern, bilinguals did not show any preference for either of their two native languages.

The possible difference regarding the capacity of bilinguals to discriminate phonetically similar languages was discarded with latter experiments (Bosch and Sebastián-Gallés, 2001). 4-month-old bilingual babies were compared to Catalan and Spanish monolingual toddlers of the same age by using the familiarization-preference procedure[8]. Monolinguals were familiarized with their native language and subsequently presented with an unfamiliarized one. Bilinguals were familiarized with just one of their native languages and subsequently presented with an unfamiliarized one. Both monolingual and bilingual infants performed longer listening times for the unfamiliarized language, revealing that all toddlers were able to discern between the two languages successfully. These results challenged the hypothesis that bilinguals would have a delay in this ability due to diverse exposure. However, bilinguals still tended to take longer than monolinguals to orient to a native language when faced with a native and a foreign one.

There are latter investigations that provide further evidence on the capacity of bilingual infants to distinguish different languages (Byers-Heinlein, Burns, & Werker, 2010) and their ability to differentiate phonologically close languages (Molnar, Gervain, & Carreiras, 2013). More modern research studied the capacity of bilinguals to discriminate both phonologically dissimilar and similar languages by using electrophysiological monitoring (Nacar Garcia et al., 2018). Their results were consistent with the reviewed above studies, suggesting that bilinguals distinguish their native languages from rhythmically dissimilar, but also from rhythmically similar languages. Nevertheless, consistent with the findings from Bosch and Sebastián-Gallés (1997) with behavioral data, this study also shows differences in the

[7]The facial reactions and movements of the infant are observed through a monitor. It is assumed that the baby will react faster to familiar stimuli when she/he is able to differentiate between two stimuli.
[8]Infants are habituated to stimuli from one language and subsequently tested on stimuli from a different one. Novel reactions, such as longer latencies, are taken as evidence for discrimination.

electrophysiological reactions between monolinguals and bilinguals, suggesting that the processes underlying language discrimination are not identical.

2.2.2. Establishment of the phonetic inventory

Bosch and Sebastián-Gallés carried out a series of experiments testing vowel discrimination among Catalan and Spanish monolinguals and bilinguals (2003a, 2005, 2009). Infants ranging from 4 to 12 months were tested using a modified version of the familiarization-preference procedure (Jusczyk & Aslin, 1995). The results for all language-specific vowels /e-ɛ/ (2003a) and language-common vowels /e-u/ vs. /o-u/ (2005, 2009) contrasts showed an unexpected pattern in the discrimination capacity in bilinguals. While, as expected, the ability of Spanish monolinguals to discriminate the non-native contrast had deteriorated by 8 months, bilinguals seemed to also suffer a decline in their perception capacity, being Catalan monolinguals the only ones that succeeded at discrimination at this age.

However, when data collected from 12-month-olds was included (Sebastián-Gallés & Bosch, 2009), both Catalan monolinguals and Catalan-Spanish bilinguals were able to discern between the language-specific vowel contrast. These results suggested that bilinguals lose and regain their ability to distinguish phonetic contrasts that are meaningful in only one of their native languages. The latter experiments regarding acoustically distant /e-u/ vs. acoustically close /o-u/ vowels common in both Spanish and Catalan replicated the bilingual-specific pattern for language-common acoustically close vowels. While both Catalan and Spanish monolinguals were able to distinguish the contrasts at all ages, bilinguals seemed to lose and regain the capacity even when the phonetic contrast belonged to both of their languages. Nevertheless, other research did not reproduce this bilingual-specific pattern development in the perceptual ability. According to the studies by Burns et al. (2007) and Sundara, Polka and Molnar (2008), French-English bilingual infants were able to discern the presented consonant contrasts at all ages.

Albareda-Castellot, Pons and Sebastián-Gallés (2011) noted that the procedure used could be underestimating the phonological discrimination capacity of bilinguals in some instances. Firstly, to test the functionality of an adaptation of the Anticipatory Eye Movements paradigm[9] (McMurray & Aslin, 2004), an experiment with the /e-u/ contrast, a language-

[9]This method assesses the discrimination capacity based on the anticipation of visual stimuli due to auditory stimuli. When the infant is able to predict the visual stimuli correctly, her or his capacity to discriminate the auditory stimuli will be evidenced.

common distant vowel contrast, whose discrimination could be observed by the former method, was conducted. Secondly, the /e-ɛ/ contrast, a Catalan-specific close vowel contrast, was tested in order to observe whether the specific pattern in bilinguals was replicated or the new procedure would unmask the ability of bilinguals at 8 months.

As expected, the results from the first experiment showed that all participants acquired the anticipation based on their capacity to discriminate between the /e-u/ vowels, proving the effectivity of the method. According to the second experiment, both Catalan monolinguals and Spanish-Catalan bilinguals acquired the anticipation based on their ability to discern the /e-ɛ/ vowels at all ages, providing evidence that the former procedure was not sensitive enough in certain conditions and that bilinguals do not suffer from a temporal failure of their discrimination capacity. Albareda-Castellot, Pons and Sebastián-Gallés (2011) attributed these divergences to the method used to measure the existence of discrimination of contrasting sounds, in particular to the familiarization-preference procedure. This measure is indirect, as it is based on attention recovery times and seemed to ignore capabilities that the Anticipatory Eye Movements paradigm detects. Furthermore, Albareda-Castellot, Pons and Sebastián-Gallés (2011) stated that this factor would not be the only reason why the discrimination capacity of bilingual infants has been misinterpreted in previous experiments, but that the ratio of cognates shared by the languages is a determinant factor for the differences found in discrimination patterns. These issues regarding earlier experimentation will be discussed in a following section.

2.2.3. Audiovisual perception of speech

Regarding the audiovisual perception of speech, Weikum et al. (2007) conducted one of the earliest studies with bilingual infants. English and French bilingual and monolingual toddlers were tested on their capacity to discriminate between these languages. It was found that babies (both monolingual and bilingual) were able to discern the familiar from the foreign language through visual information alone. Weikum et al. (2007) also showed that bilinguals maintained the discrimination capacity longer than their monolingual peers (after 8 months). Another study by Sebastián-Gallés et al. (2012) replicated the experiment with Catalan and Spanish bilinguals. According to their results, bilingual infants could distinguish English from French, while monolingual toddlers could not. These findings suggested that bilingual infants rely on the extralinguistic cues more than monolinguals due to the challenging input to which they are exposed. Further research (Birulés et al., 2019) noted that this sensitivity prevails

more in bilinguals whose languages belong to the same rhythmic category than when the languages are rhythmically different.

Pons, Bosch & Lewkowicz (2015) aimed to determine whether monolingual shifts in attention exposed in the previous section about monolinguals occurred in the same manner in bilinguals. For this purpose, the gaze of 4, 8 and 12-month-olds was traced, Catalan and Spanish monolinguals and bilinguals watched and listened to recordings of sequences in both languages. The results of the experiment with monolingual participants reproduced the findings from Lewkowicz and Hansen-Tift (2012). Regardless of the language, 4-month-old monolingual infants spent more time looking at the eyes, while 8-month-olds shifted their attention to the mouth. At 12 months, monolinguals draw their attention equally to both the eyes and the mouth of the speaker when she spoke the native language, but spent more time looking at the mouth when hearing the non-native language. Nevertheless, a different selective attention pattern was found for bilinguals. 4-month-olds draw their attention equally to both the eyes and the mouth of the talker regardless of the language, while 8- and 12-month-olds spent more time looking at the speaker's mouth for both languages.

2.3. Issues regarding experimentation

This section will review in depth the issues noted by Albareda-Castellot, Pons and Sebastián-Gallés (2011) regarding previous experiments: the head-turn preference procedure vs. the Anticipatory Eye Movements paradigm, the ratio of cognates and the nature of the speech sounds put to the test.

2.3.1. Head-turn preference procedure vs. Anticipatory Eye Movements paradigm

The head-turn preference procedure (HTPP) has a long history and an extensive amount of variations. This method provides an indication of how interesting different stimuli are to infants. Interest is measured by recording the amount of time that babies attend to the stimuli presented. By comparing infants' preference among two or more stimuli types, researchers can determine whether babies are sensitive to the relevant properties that differentiate the stimuli.

In order to conduct the HTPP, participants (usually infants between 4 and 12 months old) sit on the lap of a caretaker in a soundproof booth (see Figure 1). The toddler and the adult are positioned in front of a panel with a green bulb between two panels with a red bulb at both sides. Behind each side panel, there is a loudspeaker and behind the center panel, there

is a video camera, a computer terminal and a set of controls, that allow the examiner to turn the side and center lights on and off and record the direction and duration of the infant's head turns. Both the examiner and the adult listen to music through headphones, so they are blind to the experimental conditions. This procedure is not used with newborns younger than 4 months as it requires them to correctly locate lateral sound sources and turn their heads toward the sides (Bosch & Sebastián-Gallés, 1997). However, it has been successfully used with babies as old as 18 months old, allowing researchers to explore infant sensitivity to more advanced aspects of language, such as syntax (Santelmann & Jusczyk, 1998).

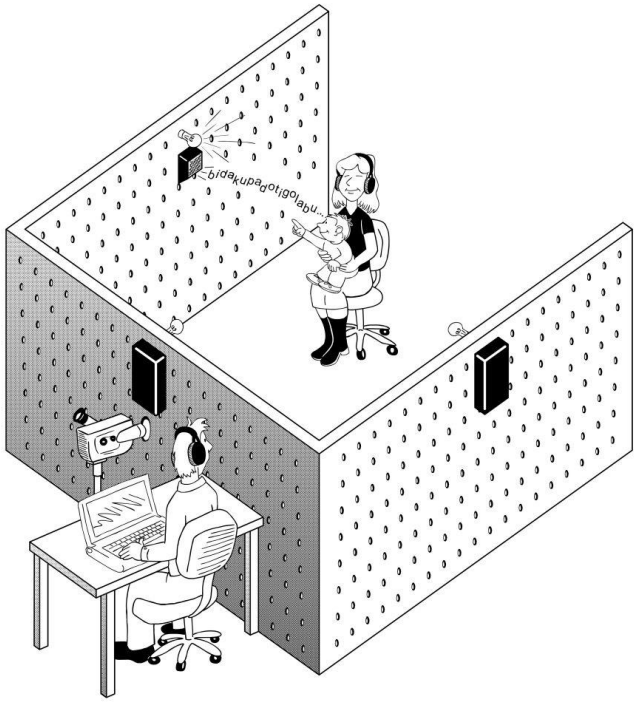

Figure 1. Layout of the testing booth by Mike Almodovar. From Brooks and Kempe (2012).

The HTPP is also known as the preference-familiarization procedure (Jusczyk & Aslin, 1995). The operation behind the methods follows the same logic, although there may be variations in the colors of the light bulbs, the amount of time the light is on in relation to the auditory stimuli, the duration of sounds presented and the amount of time they are presented (repeatedly). According to Bosch and Sebastián-Gallés (2003a: 228), the paradigm "derives from the head-turn preference procedure, but it includes an extended familiarization phase". In their study, monitors with images that appeared and disappeared on different screen locations were used instead of colored light bulbs.

All trials begin with the flashing of the green light so that the infant turns toward the center panel. Once she or he has done so, one of the red lights will flash and the auditory stimuli will be played. After a few seconds, the infant will become bored and look away. Then, the green light will flash again and a new trial will start. After a familiarization phase with sound A, the test phase will begin with sounds A and B. If the infant spends prolonged periods of time looking toward the new stimuli, it will indicate a state of surprise which in turn will reflect that she or he distinguishes between the two or more speech sounds.

It must be noted that the variations of this method, although having similar operations, may assess discrimination differently. For instance, in language discrimination experiments, extended periods of attention indicate discrimination due to the recognition and preference for native or native-like speech (Gerken, Jusczyk & Mandel, 1994). Even so, the ability to discriminate sounds is generally evidenced by different durations in the attention time for different types of stimuli.

On the other hand, the Anticipatory Eye Movements paradigm (AEM; McMurray & Aslin, 2004) assesses the discrimination capacity based on the anticipation of visual stimuli due to auditory stimuli. Different from the previous procedure, that measures attention in order to determine whether a speech sound is recognized or surprising, the AEM paradigm is a direct method. When the infant is able to predict the visual stimuli correctly, her or his capacity to discriminate the auditory stimuli will be evidenced.

During the training phase, participants learn to anticipate the location of visual stimuli at one of two spatial locations (right or left) based on the identity of two auditory stimuli. As in the HTPP, infants sit on the lap of a caretaker in a soundproof booth. The toddler and the adult, who must wear headphones to avoid manipulation, are positioned in front of a monitor, two loudspeakers (set behind the monitor) and an eye-tracking device controlled by an examiner.

> "The eye-tracker consists of a small infrared camera that captures an image of the eye and locates the pupil and corneal reflection. This camera is mounted on servomotors that are controlled by both an optical tracking algorithm and an external magnetic head-tracker [...] that locates the eye when optical tracking fails." (McMurray & Aslin, 2004: 9-10)

Each trial starts with a contracting and expanding colorful image placed at the center of the screen. Once the eye tracker (previously calibrated) detects the infant's gaze on it, the image disappears and reappears on one of the upper sides of the screen while one of two sounds is

played. After the training phase, infants are presented with a series of trials in which untrained stimuli serve as the cue to the AEM paradigm. In the adaptation from Albareda-Castellot, Pons and Sebastián-Gallés (2011), the visual stimuli (a cartoon of Sesame Street's Elmo) appeared in the lower central part of the screen below a blue T-shaped occluder (see Figure 2). Instead of disappearing and reappearing, the cartoon hid behind the occluder (moving upwards) and came out from one of the two edges (moving right or left).

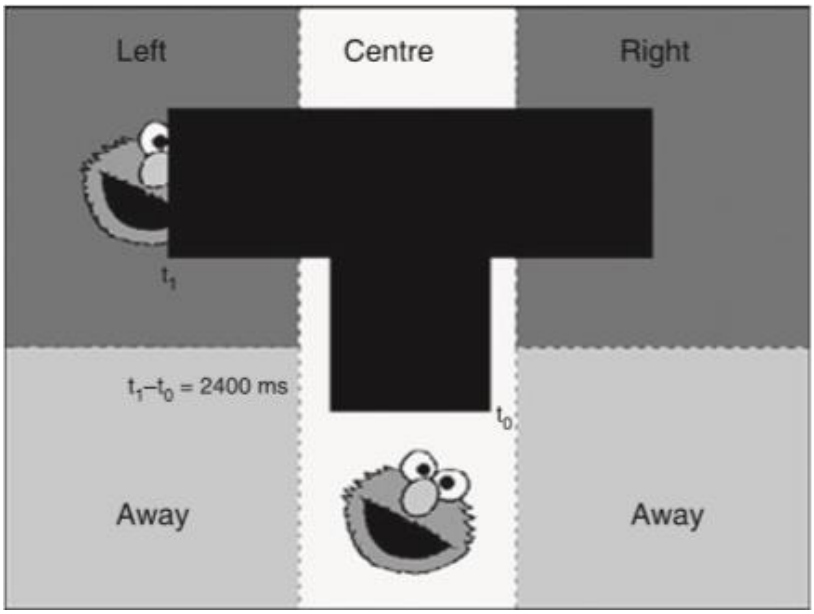

Figure 2. Visual stimuli and screen locations. From Albareda-Castellot, Pons and Sebastián-Gallés (2011).

Trials are considered correct anticipations each time the infant spends more time looking at the correct side (right or left) than the wrong side (right or left). Looks to the center or away locations (lower right and lower left) are discarded. At first, participants are expected to look randomly across the screen in search of the cartoon. After reiterated exposure to two different cueing stimuli, their ability to predict where the visual stimuli will come from based on the auditory stimuli will improve. This procedure has been successfully used with babies between 5 and 8 months of age (McMurray & Aslin, 2004; Albareda-Castellot, Pons & Sebastián-Gallés, 2011). However, eye-tracking techniques are widely used in both children and adults in different research fields, such as medicine, psychology, marketing, etc.

2.3.2. Ratio of cognates

The word cognate derives from the Latin noun *cognatus*, formed by *co-* (with) and *-gnatus* (born), originally meaning "related by blood". According to *A Dictionary of linguistics and phonetics* by Crystal (2016), a cognate is "a language or a linguistic form which is historically

derived from the same source as another language/form". Therefore, Spanish and Catalan are cognate languages, since both originated from Latin, and words like *madre* and *mare* are cognate words (or simply cognates), as both of them etymologically derived from the Latin form *mater*.

In their study, Albareda-Castellot, Pons and Sebastián-Gallés (2011: 1) suggested that "the high ratio of cognates in Spanish and Catalan may underlie the reason why bilinguals failed to discriminate the native vowels when assessed through the familiarization-preference procedure". The explanation behind this hypothesis is that vowel changes would not surprise Spanish and Catalan bilinguals as they are exposed to a large number of words that fundamentally diverge in their vowels and that, as an adaptive strategy to learn new words from both languages, they would focus their attention on consonants rather than vowels. In order to exemplify this, the crosslinguistic variations of *chocolate* were presented. This word, originary from the Nahuatl[10] language, *xocoatl* from *xoco* (bitter) and *atl* (water), is an almost worldwide cognate. Every Spanish and Catalan bilingual will be faced with word pairs like *chocolate* /ʧokolate/ and *xocolata* /ʃukułatə/, whose main phonological differences arise on their vowel sounds.

Along these lines, Mora and Nadeu (2012) observed that not only the amount of exposure to each language but also the use of cognate words were influential factors in their study. The ability to perceive and discriminate the /ɛ-e/ contrast was assessed among Spanish and Catalan bilingual adults with different degrees of exposure to each language. Both accuracy (% correct) and latencies (reaction times) were measured to assess the capacity. According to their results, participants with greater exposure to Spanish discriminated slower the Catalan language-specific contrast. More intriguingly, both groups articulated lower and fronter /ɛ/ vowels in cognate words (where Spanish has /e/), such as *maleta* /maˈlɛtə/vs. *maleta* /maˈleta/, than in non-cognate words, such as *galleda* /gaˈʎɛdə/ vs. *cubo* /ˈkubo/.

Furthermore, Ramon-Casas and Bosch (2010) previously found that toddlers were able to detect vowel mispronunciations in non-cognate words but not in cognate words. 24-month-olds from different linguistic environments (Catalan monolingual vs. Catalan and Spanish bilingual) were tested on their ability to detect vowel changes (/ɛ/ vs. /e/) among different types of words (cognate vs. non-cognate). Their results were compared to the ones achieved

[10]Classical Nahuatl was the language of the Aztec empire, used as a lingua franca in Mesoamerica from the 7th century AD until the Spanish conquest in the 16th century.

by Ramon-Casas et al. (2009) from a similar experiment conducted only with cognate words. In this study, bilinguals were not able to detect the vowel changes as monolinguals. The analysis indicated a relevant interaction between the type of words (cognates vs. non-cognates) and the experimental group (monolinguals vs. bilinguals). When the capacity to detect a vowel mispronunciation was tested on non-cognates, Catalan and Spanish bilinguals did not differ from Catalan monolinguals.

Although cognates appear to be a disadvantage when discerning specific phonetic contrasts, there is a significant number of studies on the role of cognates that discuss a facilitatory effect on speech processing and speech production at a lexical level, suggesting that bilinguals produce and recognize cognates faster than non-cognates (Sherkina-Lieber, 2004; Costa, Santesteban & Caño, 2005; Sheng et al., 2016).

2.3.3. Nature of the input

Consonants and vowels are the two main groups in which pulmonic speech sounds are classified. While vowels are produced by expelling the airflow without disruption, consonants are realized by partially obstructing the airflow through the articulatory organs. These stimuli of different nature have led to different results in the experimentation on speech perception in both children and adults.

As it has already been reviewed in previous sections, the categorical perception of consonants settles during the second half of the first year of life. Progressively, infants are no longer able to differentiate between any phonetic contrast and an inventory with the phonologically relevant categories of the native language is established. This development has been proved among different phonetic variations, such as the place of articulation, as in Werker and Tees (1984) by comparing bilabial vs. alveolar and velar vs. uvular; the manner of articulation, as in Tsao, Liu and Kuhl (2006) by comparing affricate vs. fricative; and the voicing, as in Burns et al. (2007) by comparing voiced vs. voiceless speech sounds.

On the other hand, research on discrimination between native and non-native vowels presents a less homogeneous picture among both infants and adults. Even though there are investigations on vowel discrimination (Kuhl, 1983; Marean, Werner & Kuhl, 1992), fewer studies managed to show the entire settling pattern, in which discrimination of native vowels increases and discrimination of non-native vowels decreases (Albareda-Castellot, Pons & Sebastián-Gallés, 2011). Although former misleading results, it was eventually found that

there is a similar settling pattern for (at least some) vowel contrasts. Furthermore, in their meta-analysis, Tsuji and Cristia (2014) concluded that infants' capacity to discriminate between native and non-native vowels follows different directions from the age of 6 months indeed.

However, it is still challenging to observe the ability to discriminate between specific contrasts, especially in the case of vowels. One of the main complications surrounding the study of vowel perception is the directional asymmetry found in discrimination tasks. Even within the same group tested on the same stimuli, results vary according to the order in which the stimuli are presented. Some of the latest studies in this matter (Masapollo, Polka & Molnar, 2017; Zhao et al., 2019) observed that listeners are universally biased toward focal (with prominent spectral peaks) vowels independently of language-specific prototype (from the native language) categorization, as it was formerly considered.

Furthermore, it seems that discrimination between vowels is not observed in the same manner as in consonants, because it does not occur in the same fashion. A study conducted by Altmann et al. (2014), evidenced the differences between consonant and vowel sounds on categorical perception. In their investigation, both behavioral and neurophysiological responses were measured to test the discrimination capacity. On the one hand, the behavioral results showed that vowels led to more mediocre categorical effects than consonants. On the other hand, the neurophysiological results showed an evident categorical effect for consonants, but not for vowels.

Within the consonants group, fricatives seem to be more difficult to discriminate by both infants and adults as they are acquired later than the first year of life. In pioneering research on the acquisition of fricatives in English, it was found that the tested infants were able to discriminate between /s/ and /v/ and between /s/ and /f/, but not between the voiceless /s/ and voiced /z/ sibilants (Eilers & Minifie, 1975). Eilers, Wilson and Moore (1977) conducted a similar experiment through the visually reinforced infant speech discrimination paradigm[11] instead of the high-amplitude sucking procedure. In this study, it was shown that 3-month-olds could discern between [as] and [az], but not between [sa] and [za]. However, at 6 months, the second contrast was already distinguishable, suggesting that the position of

[11]Infants are trained to turn their head when a change in stimulus is perceived by rewarding them with an automated toy every time a switch in sound is successfully detected.

the contrast is a determining variable that, in some cases, requires more experience in order to achieve discrimination.

Bosch and Sebastián-Gallés (2003b) brought bilingualism into the fricative acquisition picture with an investigation with Catalan and Spanish monolinguals and bilinguals. The voicing contrast /s/ vs. /z/, that exists in Catalan but not in Spanish was tested using the preference-familiarization procedure. No participants showed discrimination for word-initial [siði] and [ziði] and word-medial [disi] and [dizi] positions, while all three groups did for word-initial [sit] and [zit]. On the other hand, a more recent bilingual study regarding fricatives and affricates also found that this group of consonants is more difficult to discriminate than others (Larraza, Molnar & Samuel, 2020). Basque and Spanish monolingual and bilingual infants ranging from 6 to 7 months and 11 to 12 months were tested through the visual habituation procedure[12] on their ability to discern between apical /s̻/ and laminal /ʂ/ alveolar fricatives and between the palatal /tʃ/ and alveolar (laminal) /tʂ/ affricates. All participants carried out the task poorly, and no differences due to language exposure nor age were observed.

Once again, it must be remarked that the methodology may not have been adequate. The first investigation assessed discrimination through the preference-familiarization procedure and the second through the visual habituation procedure. Albareda-Castellot, Pons and Sebastián-Gallés (2011) evidenced in their study that this type of methods (that rely on surprise to assess discrimination are not the most suitable for some speech contrasts. In their investigation, Albareda-Castellot, Pons and Sebastián-Gallés (2011) also suggested that the comparisons between the studies, not only did not consider the differences between the language pairs (ratio of cognates) but also did not take into account the differences between the stimuli tested (nature of the speech sound). As reviewed in this section and previous ones, it is impractical to draw conclusions based on the comparison across studies that differ in these matters.

2.4. The language pairs

For the psycholinguistics field, it is essential to know the language(s) through which the language capacity is to be studied. What is the phonology of this language like? And its

[12]Infants are presented auditory stimuli as long as they look at visual stimuli. Habituation phase is prolonged until the infant's looking times lower till a determined criterion. In the testing phase, novel auditory stimuli are presented and longer latencies evidence discrimination.

syntax? Or its morphology? Is it similar to the other languages that coexist in the territory? Nevertheless, the different use and prestige of languages in society can also be a determining factor. Do the speakers use one language more than the other? Do they have equal linguistic competencies for both of them?

2.4.1. Basque Country

Basque is an agglutinative, ergative isolated language. It is spoken in the Euskal Herria, a historical territory that shares the Basque language and culture. It comprises areas from both sides of the Pyrenees, Spain (Araba, Bizkaia, Gipuzkoa and Nafarroa Garaia) and France (Lapurdi, Zuberoa and Nafarroa Beherea). The Statute of Autonomy for the Basque Country establishes that Basque is the official language in the Basque Country, as Spanish, which is the official language of the Spanish state, and that all inhabitants have the right to know and use both languages.

According to data compiled by Eustat (Basque Statistics Office) for the Population and Housing Statistics from 2016, 28% of the population (age range was not specified) resident in the autonomous community of the Basque Country was born elsewhere. Within this group, 68% of this "non-Basque" population was born in the Spanish territory, and 32% was born in a foreign land (2018). Although the majority of the inhabitants are born in the Basque Country, and this population has been defined as sedentary by the same Eustat study, the population's knowledge of the co-official languages is far from being equal. According to the *VI Encuesta Sociolingüística* (Linguistic Policy Council, 2016), only 33.9 % of the population (over 15 years old) living in the Basque Country speaks Basque, 19.1 % is passive Basque-speaking, meaning that they understand but do not speak the language, and 47 % is monolingual in Spanish. It must be noted that there are also differences between the regions: half of the population of Gipuzkoa is Basque-speaking (50.6%), while in Bizkaia just over a quarter (27.6%) and in Araba less than one fifth (19,2 %) of the inhabitants knows Basque.

Regarding the daily use of the language, according to the same report (Linguistic Policy Council, 2016), it is estimated that more than half of the population (63.3%) always speaks Spanish, French or another language, while only a little more than one-tenth (13.4%) speaks more (but not always) in Basque than in Spanish, French or another language. On the other hand, 7.1% consider that they speak Basque and Spanish, or French or another to the same extent. The remaining percentage (16.2%) encompasses people who speak more Spanish, French or another language than Basque and people who claimed to speak very few Basque.

In general, the use of Basque has increased in all areas of use: at home, with colleagues, with friends and in the formal sphere. The most significant increment has taken place in the formal setting, especially in municipal and health services. In 2016, 20.6% of the inhabitants of the whole of the Basque territory used Basque as much or more than Spanish in municipal services and 15.6% in health services, while they were 11.7% and 6.9%, respectively, in 1991.

2.4.2. Catalonia

Catalan is a fusional, nominative romance language. It is spoken in the Països Catalans, a cultural territory that shares the Catalan language. It comprises the Spanish autonomous communities of Catalonia, the Balearic Islands and the Community of Valencia; the Pyrenean microstate of Andorra; the area of Catalan-speaking territories of eastern Aragon bordering Catalonia (La Francha, Spain); the city of Alghero (Sardinia, Italy); the historical territory of Roussillon (Eastern Pyrenees, France) and the small territory of El Carche (Murcia, Spain). The Statute of Autonomy for Catalonia establishes that Catalan is the official language of Catalonia, as Spanish, which is the official language of the Spanish state, and that all inhabitants have the right to know and use both languages.

According to the *Enquesta d'usos lingüístics de la població* carried out by Idescat (Statistical Institute of Catalonia), only 60% of the Catalan population (over 15 years old) was born in Catalonia in 2018. Within this group, more than half of this percentage represents people whose parents (at least one of them) were born outside the autonomous community. The remaining 40% of the population is equally divided into people born within the rest of Spain and people born in foreign countries (2019). Before presenting the use of the co-official languages, it should be borne in mind that there are differences between the knowledge of the inhabitants with respect to the two languages. Language knowledge was divided into four competencies: speaking, writing and written and oral comprehension. According to their surveys, it is estimated that 99.8% of the population understands Spanish, 99.5% is able to speak it, 98.5% knows how to read it and 97.6% to write it, while 94.4% understands Catalan, 81.2% is able to speak it, 85.5% knows how to read it and 65.3% to write it. Differently from the Basque Country, the results were not as dispersed across the regions.

According to the same report (Idescat, 2019), it is estimated that 93.2 % of the population uses Spanish on a daily basis, while 73.4% uses (or also uses) Catalan. There is relative diversity in use regarding the situation type. In formal situations, the percentage of inhabitants who consider that they use only Catalan or more Catalan than Spanish range

from 47.5% (local administration) to 31.6% (state administration), while the use of only Spanish or more Spanish than Catalan varies between 47.9% (state administration) and 33.8% (administration of the Generalitat). Other formal environments that range between those proportions are bank, medical and businesses staff. The percentages of balanced use among those vary between 7.8% to 18.7%. In informal situations, 36.3% of the students and 30.4% of the workers claim to speak only Catalan or more Catalan than Spanish, while 32.8% (students) and 39.2% (workers) consider that they speak only Spanish or more Spanish than Catalan. 19.1% of the students and 20.2% of the workers estimate that their language use is balanced. The proportions for at home and with friends contexts are 32.8% and 29.9% respectively for only Catalan or more Catalan than Spanish, while 47.8% and 42% accordingly for only Spanish or more Spanish than Catalan. 7.8% (at home) and 18.3% (with friends) of the population estimate that their language use in these environments is balanced.

2.4.3. Educational systems

Spain delegates the administration of education and language policy to the Autonomous Communities. Article 15, Law 10/1982, dated 24th November, from the *Boletín Oficial del País Vasco* recognizes the right of all students to receive instruction both in Basque and in Spanish at the various educational levels. Accordingly, article 20, Law 1/1998, dated 7th January, from the *Diari Oficial de la Generalitat de Catalunya,* states that schools at any level must have Catalan as the usual vehicular language in teaching and administration activities, both internally and externally.

The Basque Country has a three-track educational system: Model A, in which Spanish is the vehicular language (and Basque is studied as any other subject); model B, in which Basque and Spanish are used equally among the different subjects; and model D, in which Basque is the vehicular language (and Spanish is studied as any other subject). The Eustat (2019) estimated that, in the academic year 2019/20, 256.909 students would be matriculated for model D, while 58.722 and 68.665 would be matriculated for models A and B accordingly. The data are based on both public and private educational centers without taking into consideration university institutions. On the other hand, in Catalonia, the vehicular language of every public education center is Catalan, while Spanish is taught a certain number of hours per week depending on the school grade.

2.4.4. Linguistic profiles

All these groups have an infinite number of different linguistic profiles. In the Catalan context, there will be people who have mastered both languages since childhood, children who have not had any contact with Catalan before going to school, but who have had contact with Spanish. Others may not have established contact with any of the two official languages before entering the educational environment. Similar situations can be predicted from the Basque context. When adding the three-track educational system, the equation becomes even more complicated. According to the data earlier presented, most children will not get in contact with the Basque language until they start school and, once they do, their exposure to the language will vary from model to model. It is worth noting that, while the situation surrounding Catalan is rather stable, the sociolinguistic landscape of Basque is experiencing a period of transition. Only 30 years ago, model A represented 66.6% of the matriculations, while model D represented only 18.9%. To this day, the Basque community is committed to recovering the use of its own language and these numbers have completely turned (see Figure 3), 15.8% of the students are enrolled in model A, while there are 65.7% for model D (Eustat, 2020).

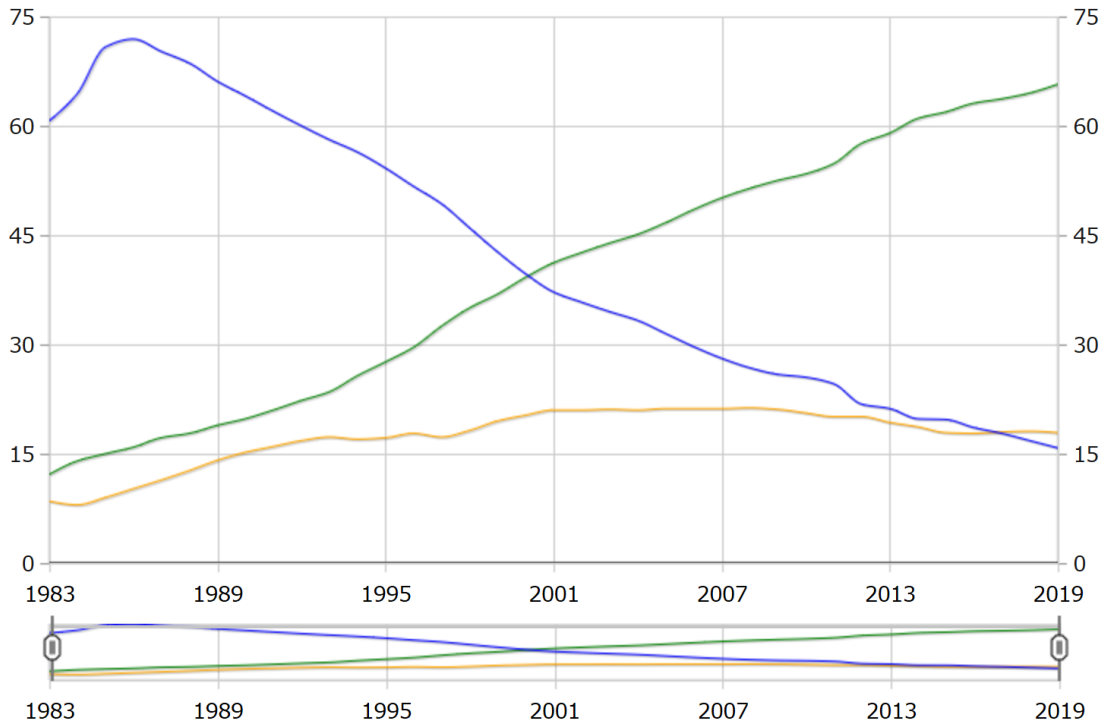

Figure 3. Percentage of students matriculated in non-university general education by language model in the Basque Country. Blue represents model A, green, model D and orange, model B. From Eustat (2020).

3. PROPOSAL

3.1. General description

This proposal describes an experiment based on the two-alternative forced-choice Anticipatory Eye Movements paradigm (McMurray & Aslin, 2004). It aims to test the ability to discriminate speech sounds in two groups of bilingual infants at different ages to establish their evolution. The objective of this proposal is to determine whether the ratio of cognates shared by the bilingual participants' languages has any measurable consequences concerning their development in the acquisition of phonological categories. In order to do so, the discrimination results from Catalan and Spanish and Basque and Spanish bilinguals will be analyzed, as these two language pairs differ noticeably in their ratio of cognates. The speech contrasts will be two articulations of the alveolar /s/ and the postalveolar (or pre-palatal) /ʃ/ fricative pulmonic consonants, since the first sound is a distinctive phoneme in all three languages (Spanish, Basque and Catalan), while the second one is only contrastive in one of the languages from each group (Catalan and Basque). Spanish monolinguals will also be tested as a control group to assess whether this specific contrast becomes invisible when it does not belong to the phonetic inventory of the native language.

3.2. Empirical contribution

This proposal is based on and inspired by the work of Albareda-Castellot, Pons and Sebastián-Gallés (2011). Nevertheless, it encompasses a different goal. Although there are already a number of studies that seek to compare the development of language acquisition between monolinguals and bilinguals, the present proposal will attempt to compare this categorical perception process between bilinguals that differ in the ratio of cognates shared by their languages. For this purpose, not only bilinguals in Catalan and Spanish will be tested, but also bilinguals in Basque and Spanish. The auditory stimuli will also be distinct. While the research by Albareda-Castellot, Pons and Sebastián-Gallés (2011) and the previous ones by Bosch and Sebastián-Gallés (2003a, 2005, 2009) were conducted with vowels, a consonantal contrast has been selected for this study design. Furthermore, the participants are considerably older, and a longitudinal study has been preferred in order to reduce sampling errors.

3.3. Participants

The participants will be two groups with the same infants tested at 1, 2 and 3 years old as fricative categories seem to settle later than the first year of life (Larraza, Molnar & Samuel, 2020). The participants will be recruited in the central regions of the Basque Country and Catalonia, Gipuzkoa and Barcelona. Thus, the possible exposure to the standard (used in the media), will not be detrimental to our purposes, as the "unification" of these languages was primarily based on the central dialects. All infants will be full-term with no reported health problems. In case the participant is diagnosed with a health-related issue, she or he will be excluded from the sample.

It would be ideal that the participants recruited will be of a variety of ages and with a balanced bilingual profile. For this reason, the parents or caretakers of the infants will fill out a detailed questionnaire about the language exposure to which the participant is subjected during an ordinary week before each session. To be considered as a balanced bilingual in this study, the percentage of daily exposure to the native languages (Spanish and Catalan or Spanish and Basque) will have to be distributed ranging from 50%–50% to 60%–40%. In case that the infants' linguistic environment is no longer considered balanced due to changes in language exposure, the participant will be discarded from the sample. During the first phase, there will be 90 infants around 1 year old, 45 balanced Catalan and Spanish bilinguals (CS) and 45 balanced Basque and Spanish bilinguals (BS). The following phases will be conducted when the same participants are 2 and 3 years old, in order to achieve a longitudinal study. Parental consent will be signed prior to each experiment during the course of the investigation.

The initial sample will be rather large to compensate for possible reductions of the group due to crying, fussiness, interference of the caretaker, etc. Furthermore, infants who direct more than 75% of their gazes throughout the experiment to one of the sides of the screen while the visual cue is hidden will be excluded from the sample. In order to be included, participants will be required to direct their gaze to at least one of the reinforcers' reappearance locations in a minimum of 9 trials per block. These criteria are based on the procedure by Albareda-Castellot, Pons and Sebastián-Gallés (2011) and designed to ensure that the infant is involved in the task and not biased toward one side.

3.4. Stimuli and variability

The auditory stimuli will consist of 24 natural samples of the monosyllabic CV non-words [sa] and [ʃa] to test the ability to discriminate between the alveolar and postalveolar (or pre-palatal) fricatives /s/ and /ʃ/. The first speech sound is a distinctive phoneme in Spanish, Basque and Catalan, while the second is only contrastive in the Basque and Catalan languages. This design will allow testing the same contrast (fricatives varying on the place of articulation) among different groups of bilinguals (high ratio of cognates vs. low ratio of cognates). In order to ensure that discrimination is based on the existence of different phonological categories and not merely on acoustic cues, the stimuli will be recorded by 8 female speakers and several tokens from each will be selected. Furthermore, all tokens will be produced in infant-directed speech, as it is commonly preferred by infants (Cooper & Aslin, 1990). Infant-directed speech, also known as motherese or baby talk, is characterized by a higher pitch, extended intonation contours and exaggerated phonetic cues. The variability of the auditory stimuli will increase progressively to facilitate the familiarization process and in a semi-randomized to maintain a maximum of 3 equal trials in a row.

The visual stimuli used as a reinforcer to teach anticipation will be a colorful image of Donald Duck's face from modern Disney cartoons, while the occluder will be a black capital T shape so that it does not purposely attract attention. The screen will be invisibly divided into 5 locations. Center will be a column-shaped area wider than the vertical axis of the occluder on each side. The remaining right and left areas will be subdivided into upper and low locations. Upper sides will cover the area from the screen corner till the corner created at the corresponding side of the occluder below its horizontal axis. The remaining areas will be categorized as away locations since they are the furthest from the target positions (the edges of the horizontal axis of the occluder). The Donald Duck cartoon will be 16 cm wide and 20 cm high and the T-shaped occluder will be 60 cm wide and 42 cm high. The variability of the visual stimuli will proceed according to the auditory stimuli, in order to establish a correlation relationship between the upper locations and the speech sounds tested. There will be 8 different presentation orders, generated by crossing the sounds and the sides of reappearance of the reinforcer.

3.5. Task and measurements

The procedure will consist in a contrast discrimination task through an adaptation of the two-alternative forced-choice Anticipatory Eye Movements paradigm (McMurray and Aslin, 2004) as it has been demonstrated to be a more sensitive method than the head-turn preference procedure (Albareda-Castellot, Pons & Sebastián-Gallés, 2011). Participants will be presented a series of visual stimuli, accompanied by auditory stimuli. After a period of familiarization (Block 1), it will be possible to check if the infants are able to discriminate between the speech sounds, depending on whether they can anticipate the visual stimuli when hearing the auditory stimuli (Block 2). At first, participants are expected to look randomly across the screen in search of the cartoon. After reiterated exposure to two different cueing stimuli, it is expected that their ability to predict where the visual stimuli will come from based on the auditory stimuli will improve.

As in the adaptation of Albareda-Castellot, Pons and Sebastián-Gallés (2011), each trial will begin with the cartoon, in this case of Donald Duck, which will expand and contract below the T-shaped occluder to attract the participants' attention as far as possible from the two possible subsequent locations. Once the infant has directed her or his gaze toward the reinforcer, the image will hide under the occluder and travel at a constant velocity. After 2500 ms, it will come out from one of the upper edges (right or left) and rotate to attract the attention of the participant. During this process, the participant will hear the auditory stimuli assigned to the given side 5 times. There will be 46 trials and; once every 4 trials, the infant will be presented with one of three short animations of colored bubbles with soft background music to re-involve her or him in the task. Participants not reaching a minimum of 18 valid trials, due to side bias, crying, fussiness or interference of the caretaker will be excluded from the experiment.

3.6. Conditions and materials

The experiment will take place in a soundproof booth. The infant will remain seated on the lap of a trusted caretaker, who will listen to music using headphones to prevent any possible manipulation of the participant's responses. They will be positioned in front of a screen at least 50 cm apart from the participants. The animations will be displayed by a projector on the screen, and the audio tracks will be played by a pair of speakers set behind the screen. An experimenter from outside the room will control both devices through a monitor. The experiment will be recorded with a video camera and an eye-tracker device located under the screen.

3.7. Data analysis

Each frame (40 ms) will be annotated with the locations of infant's gaze (right, left, center or away) from the moment the cartoon starts moving until its reappearance by a professional coder. An additional coder will replicate this task for 25% of the data, and the Pearson correlation coefficient will be calculated in order to assess their agreement. As in McMurray and Aslin (2004), trials will be considered as correctly anticipated when the participant spends more time looking to the correct side of the screen (upper right or upper left) than to the incorrect one and as incorrectly anticipated when the participant spends more time looking to the wrong side of the screen (upper right or upper left) than to the correct one. Looks shorter than 80 ms and looking times to center or away locations (lower right or lower left) will not be included in the analysis. The 46 trials will be divided into two blocks: the familiarization phase (Block 1) and the testing phase (Block 2). The percentages of correct and false responses will be computed for each block (Block 1 vs. Block 2) for each bilingual group (BS vs. CS) for each of the experiments along with the longitudinal study (1 year vs. 2 years vs. 3 years). Therefore, it will be possible to determine whether there are differences between the evolution of the different bilingual groups.

3.8. Experimental predictions

In order to determine the existence of an influence due to the ratio of cognates shared by the language pair, the development of the categorical perception of the speech sounds will be compared among the different groups of bilinguals. For this purpose, repeated measures analysis of variance will be conducted with age and bilingual group. In case of acknowledging a significant difference in the anticipation capacity in Block 2 for one of the two types of bilinguals, as represented in Figure 4, we could state that the ratio of cognates of the language pair affects phonological categorization. It is expected that this difference would be reflected by a more prominent slope; in other words, a considerably higher percentage of correct predictions for one of the language combinations.

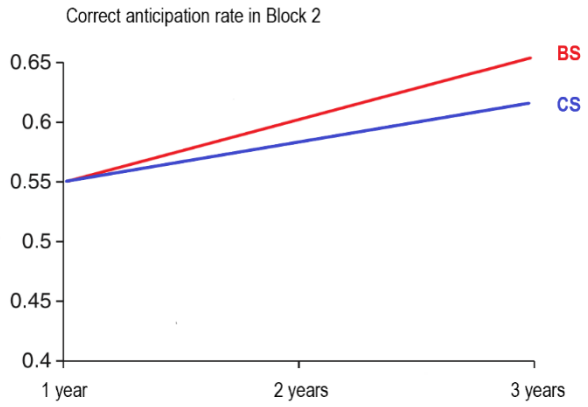

Figure 4. A possible proportion of correctly anticipated trials from Catalan and Spanish bilinguals (CS) and Basque and Spanish bilinguals (BS) when the ratio of cognates plays a role in speech acquisition.

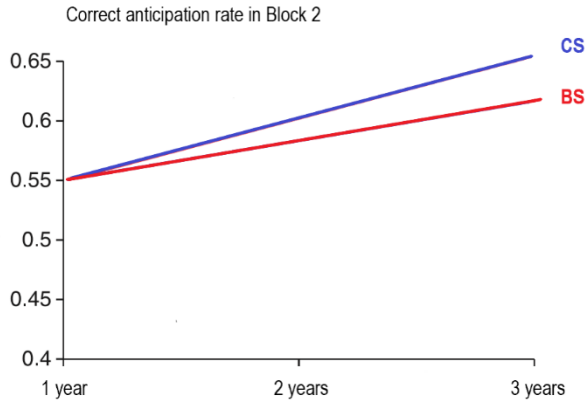

Figure 5. Another possible proportion of correctly anticipated trials from Catalan and Spanish bilinguals (CS) and Basque and Spanish bilinguals (BS) when the ratio of cognates plays a role in speech acquisition.

T-test comparisons of means against chance will be used to assess whether the difference between groups occurs by chance or the discrimination ability of one of the two language pairs is less sensitive. When the results from the different bilingual groups would not be observed (or not reach a significance), it would mean that the discrepancies regarding bilingual acquisition found in other studies have been caused by other variables. It must be noted that the results for the null hypothesis could be diverse, such as one of the groups being slightly better (Figure 5) or slightly worse (Figure 6) but not reaching the established threshold of significance of $p<.05$ or both groups achieving identical results in the task (not likely).

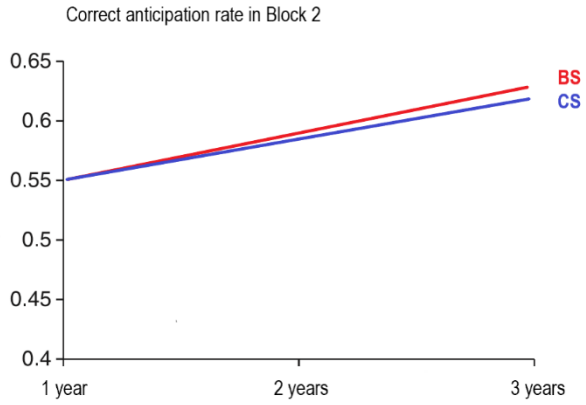

Figure 6. A possible proportion of correctly anticipated trials from Catalan and Spanish bilinguals (CS) and Basque and Spanish bilinguals (BS) when the ratio of cognates does not play a role in speech acquisition.

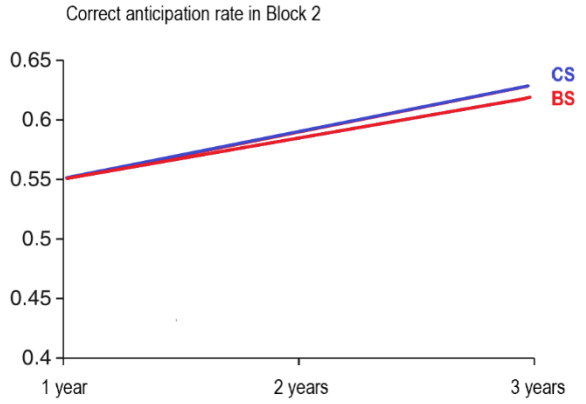

Figure 7. Another possible proportion of correctly anticipated trials from Catalan and Spanish bilinguals (CS) and Basque and Spanish bilinguals (BS) when the ratio of cognates does not play a role in speech acquisition.

3.9. Discussion

Language acquisition in bilinguals is influenced by factors that do not occur in the monolingual setting. This experiment would determine whether the ratio of cognates shared by the languages of a bilingual-to-be individual is one of these factors. If Catalan and Spanish bilinguals were observed to have an advantage over Basque and Spanish bilinguals, a facilitatory cognate effect for consonants would be suggested, as opposed to the detrimental effect for vowels discussed earlier.

Although traditionally, the cognate effect is often interpreted as a facilitator, this study could reveal the opposite scenario. In the event that Basque infants have better results in anticipating visual stimuli, it could be suggested that they have an advantage in discriminating the auditory stimuli. Conversely, Catalan children would have greater difficulty carrying out the task, indicating that, at least in the case of bilingualism in Catalan and Spanish, the presence of cognates would be detrimental.

Albareda-Castellot, Pons and Sebastián-Gallés (2011) suggested that the nature of this phenomenon would lie in the fact that part of the lexicon of a bilingual individual, whose languages share a high number of cognates, would be integrated. In addition, Mora and Nadeu (2010) found in their study that cognates influence each other across languages. However, these researchers addressed the perception and production of vowels. Results such as those described above would indicate that this disadvantage is general since the experimental proposal task seeks to observe consonant discrimination.

When no significant differences were observed between the bilingual groups, it would indicate that the cognate ratio does not influence the perceptual abilities of speech in consonant discrimination. Even so, it would be of great interest to replicate the experimental design with auditory vowel stimuli with the necessary adjustments. Unfortunately, this will not be possible with the present pair of languages.

Collaterally, it might be possible to find the temporal course of the acquisition of the categorical perception of fricatives. Several researchers have evidenced that discrimination of fricatives is later acquired and that, even in adulthood, fricative contrasts are more difficult to detect (Eilers & Minifie, 1975; Eilers, Wilson & Moore, 1977; Bosch & Sebastián-Gallés, 2003b; Larraza, Molnar & Samuel, 2020). For this reason, participants were required to perform the task annually from the first year of life to 3 years of age.

3.10. Limitations and difficulties

3.10.1. Regarding participants and their language use

It should be borne in mind that, although both communities are bilingual, the Basque Country and Catalonia differ widely in language use. Therefore, it will be more difficult to find balanced bilinguals in Gipuzkoa than in Barcelona. Furthermore, it is practically impossible to find bilinguals who receive the exact amount of input from one language as from the other. However, the young age of the participants reduces this noise to a minimum, since during the first years of life, contact with the rest of society is almost non-existent. Another difficulty that this investigation would face is that, since the participants come from different communities, not only in terms of culture and language but also geographically, the (extralinguistic) differences between the two groups could be more significant than those presented by groups that coexist in the same community, as in the case of monolingual and bilingual infants from previous studies.

3.10.2. Regarding the language pair and the speech contrast

Although Basque is an isolated and unique language, meaning that it is not related to Latin nor any other language known to date, it must be taken into account that it has adopted words from Spanish and Latin over time. Furthermore, according to Nespor, Shukla and Mehler (2011), Basque and Spanish are phonetically more similar than Catalan and Spanish. Taking these factors into consideration, Basque is still expected to diverge widely from Catalan in terms of cognates shared with Spanish. However, the fact that Basque has more in common with Spanish in relation to general phonetic characteristics could influence the proposed research toward another direction.

On the other hand, Basque and Catalan have different fricative systems. While Catalan discerns between voiced /z/ and voiceless /s/ alveolar fricatives, Basque distinguishes between apical /s̻/ and laminal /s̺/ alveolar fricatives. However, it is expected that the difference in place of articulation between alveolar /s/ and postalveolar /ʃ/ contrasts can be fairly compared. Another limitation that researchers usually take into consideration is whether one of the phonemes tested is more frequent than the other. When comparing language-specific and language-common contrasts among two language pairs, this factor becomes even more complicated. It should be measured how frequent is the language-specific /ʃ/ compared to the language-common /s/ in both Basque and Catalan.

4. CONCLUSIONS

As reviewed in the synthesis of previous studies, the research on categorical language acquisition began around the 1970s. Although the first studies were conducted with monolingual speakers, it is currently directed to bilingual samples. Alternatively, experiments with both monolingual and bilingual participants are common, as most studies attempt to compare the differences among these linguistic profiles. With the exception of the work by Bosch and Sebastián-Gallés (2003a, 2005, 2009) with Catalan-specific vowels, the study of categorical speech perception has focused primarily on consonant contrasts. According to several studies discussed above, the discrimination of vowels and consonants seems to develop in a different manner and, at all events, it is challenging to observe with close vowels.

Another aspect that has marked the course of the study in this area of psycholinguistics is the evolution of its methodologies. Newborns younger than 4 months need to be assessed through procedures suitable for this age. One of the most appropriate methods for these young participants is the high-amplitude sucking procedure, in addition to gaze monitoring. Older infants can perform tasks that require more active participation, such as the variations of the head-turn preference procedure or the Anticipatory Eye Movements paradigm, which has been proved to be more sensitive in categorical perception experiments. More recently, brain imaging and monitoring techniques have gotten attention. These are very practical, as they can reveal cues about the cognitive processes underlying language acquisition that cannot be observed through behavioral tasks.

The third problem, exclusive to the categorical speech perception in bilinguals, is the possible influence of the cognate ratio. The cognates that a bilingual's languages share can mask mispronunciations that are detected when the words in this matter come from different sources and therefore do not resemble each other. Furthermore, it has been suggested that a high ratio of cognates may cause differences in language discrimination in bilinguals. Lastly, an experimental design has been presented. The proposal aims to test this possible influence on the discrimination ability with bilinguals in Catalan and Spanish and Basque and Spanish since these languages differ considerably in their number of cognates. However, the contrast chosen had to be consonantal, as Basque does not have vowels that it does not share with Spanish. This specific contrast from Catalan and Basque is fricative, a section of the phonetic chart that has not yet been clearly observed in this area. Therefore, despite the fact that additional adaptations had to be implemented, this experiment could be of great interest.

5. REFERENCES

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.

Altmann, C. F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., & Fukuyama, H. (2014). Categorical speech perception during active discrimination of consonants and vowels. *Neuropsychologia*, *64*, 13-23.

Birulés, J., Bosch, L., Brieke, R., Pons, F., & Lewkowicz, D. J. (2019). Inside bilingualism: Language background modulates selective attention to a talker's mouth. *Developmental Science*, *22*(3), e12755.

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. (2003a). 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.

Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In *Proceedings of the 15th international conference of phonetic sciences* (pp. 1987-1990). UAB/Casual Prods.

Bosch, L., & Sebastián-Gallés, N. (2005). Developmental changes in the discrimination of vowel contrasts in bilingual infants. In *Proceedings of the 4th International Symposium on Bilingualism* (pp. 354-363). Cascadilla Press Somerville, MA.

Brooks, P. J., & Kempe, V. (2012). *Language development*. John Wiley & Sons.

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., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological science*, *21*(3), 343-348.

Chong, A. J., Vicenik, C., & Sundara, M. (2018). Intonation plays a role in language discrimination by infants. *Infancy*, *23*(6), 795-819.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584-1595.

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and language*, *94*(1), 94-103.

Crystal, D. (2011). *A dictionary of linguistics and phonetics* (Vol. 30). John Wiley & Sons.

Eilers, R. E., & Minifie, F. D. (1975). Fricative discrimination in early infancy. *Journal of Speech and Hearing Research*, *18*(1), 158-167.

Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research*, *20*(4), 766-780.

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

España. Law 1/1998, dated 7th January, Language Policy [Internet]. Boletín Oficial del Estado, 11th February 1998, no. 36, pp. 4809 to 4818. [accessed 5th May 2020]. Retrieved from https://www.boe.es/eli/es-ct/l/1998/01/07/1

España. Law 10/1982, dated 24th November, Standardization of the use of the Basque language. [Internet]. Boletín Oficial del Estado, 26th April 2012, no. 100, pp. 32138 to 32145. [accessed 5th May 2020]. Retrieved from https://www.boe.es/eli/es-pv/l/1982/11/24/10

Eustat. (2018, April 4). *El 28% de las personas residentes en la C.A. de Euskadi en 2016 había nacido fuera de ella* [Press release]. Retrieved from https://www.eustat.eus/elementos/El_28_de_las_personas_residentes_en_la_CA_de_Euskadi_en_2016_habia_nacido_fuera_de_ella/not0014789_c.html

Eustat. (2019, November 12). *Alumnado matriculado en enseñanzas de régimen general no universitarias en la C.A. de Euskadi por territorio histórico y nivel de enseñanza, según titularidad*

del centro y modelo lingüístico. Avance de datos 2019/20 [Press release]. Retrieved from https://www.eustat.eus/elementos/El_28_de_las_personas_residentes_en_la_CA_de_Euskadi_en_2016_habia_nacido_fuera_de_ella/not0014789_c.html

Eustat. (2020, March 16). *Alumnado matriculado en Enseñanzas de Régimen General no Universitarias por modelo lingüístico en la C.A. de Euskadi* [Press release]. Retrieved from https://www.eustat.eus/elementos/El_28_de_las_personas_residentes_en_la_CA_de_Euskadi_en_2016_habia_nacido_fuera_de_ella/not0014789_c.html

Garcia, L. N., Guerrero-Mosquera, C., Colomer, M., & Sebastián-Galles, N. (2018). Evoked and oscillatory EEG activity differentiates language discrimination in young monolingual and bilingual infants. *Scientific reports*, *8*(1), 2770.

Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*(3), 237-265.

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

Idescat. (2019, July 8). *Enquesta d'usos lingüístics de la población* [Press release]. Retrieved from https://govern.cat/govern/docs/2019/07/08/13/16/dd9ed981-8399-426b-8245-729a01baf8cf.pdf

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive psychology*, *29*(1), 1-23.

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, *6*(2-3), 263-285.

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, *67*(5), 713-727.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138-1141.

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

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.

Larraza, S., Molnar, M., & Samuel, A. G. (2020). Phonemic contrasts under construction? Evidence from Basque. *Infancy*, *25*(3), 304-318.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*(5), 1431-1436.

Linguistic Policy Council. (2016, October 14). *VI Encuesta Sociolingüística: Comunidad Autónoma de Euskadi* [Press release]. Retrieved from https://www.euskara.euskadi.eus/contenidos/informacion/argitalpenak/es_6092/adjuntos/VI%20_INK_SOZLG_EAE_Presentacion_publica_20161014.pdf

Marean, G. C., Werner, L. A., & Kuhl, P. K. (1992). Vowel categorization by very young infants. *Developmental Psychology*, *28*(3), 396.

Masapollo, M., Polka, L., Molnar, M., & Ménard, L. (2017). Directional asymmetries reveal a universal bias in adult vowel perception. *The Journal of the Acoustical Society of America*, *141*(4), 2857-2869.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746-748.

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

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*(3), 326-337.

Mora, J. C., & Nadeu, M. (2012). L2 effects on the perception and production of a native vowel contrast in early bilinguals. *International Journal of Bilingualism*, *16*(4), 484-500.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407-420.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1-19.

Nespor, M., Shukla, M., & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. *The Blackwell companion to phonology*, 1-13.

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

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological science*, *26*(4), 490-498.

Ramon-Casas, M., & Bosch, L. (2010). Are non-cognate words phonologically better specified than cognates in the early lexicon of bilingual children? In *Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 31-36). Sommerville, MA: Cascadilla.

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.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292.

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophys*ics, 59(3), 347-357.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105-134.

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., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, *23*(9), 994-999.

Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, *141*, 229-238.

Sherkina-Lieber, M. (2004). The cognate facilitation effect in bilingual speech processing; the case of Russian-English bilingualism. *Cahiers linguistics d'Ottawa*, *32*, 108-121.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, *259*(5538), 39-41.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of America*, *26*(2), 212-215.

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.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America*, *120*(4), 2285-2294.

Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental psychobiology*, *56*(2), 179-191.

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*(5828), 1159-1159.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 349-355.

Zhao, T. C., Masapollo, M., Polka, L., Ménard, L., & Kuhl, P. K. (2019). Effects of formant proximity and stimulus prototypicality on the neural discrimination of vowels: Evidence from the auditory frequency-following response. *Brain and language*, *194*, 77-83.