Budding bilinguals: Investigating bilingual infants’ language acquisition

Esther Schott

A Thesis
In the Department
of
Psychology

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Psychology) at
Concordia University
Montréal, Québec, Canada

June 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Esther Schott

Entitled: Budding bilinguals: Investigating bilingual infants' language acquisition

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Psychology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair

Dr. Rosemary Reilly

______________________________ External Examiner

Dr. Julien Mayor

______________________________ Examiner

Dr. Diane Poulin-Dubois

______________________________ Examiner

Dr. Kristen Dunfield

______________________________ Examiner

Dr. Linda Polka

______________________________ Thesis Supervisor

Dr. Krista Byers-Heinlein

Approved by ______________________________

Dr. Kristen Dunfield, Graduate Program Director

Month/day/year ______________________________

Dr. Pascale Sicotte, Dean

Abstract

Budding bilinguals: Investigating bilingual babies' language acquisition

Esther Schott, Ph.D.

Concordia University, 2022

Almost one in six children in Canada grow up hearing two languages. Bilingual children frequently encounter switches between their languages, and must learn words in both. My thesis tackles three central issues of early bilingual language development.

Manuscript 1 investigated whether infants can detect language switches at the level of individual words. This ability could help infants cope with rapid language switching in their language input, and prepare them, once they start speaking, to have control over which language they use. I tested bilingual and monolingual 8- to 12-month-olds' interest in single-language trials ("milk…dog") and switched-language trials ("milk…*chien*"). Neither group showed evidence of differentiating between contexts, suggesting that detection of single-word language switching is more difficult than previously assumed based on prior research investigating multi-sentence language switching.

Manuscript 2 examined whether infants can associate a person with the language that person is speaking. Some theories of early bilingualism propose that person–language associations help infants navigate their bilingual language input. I tested 5- to 18-month-olds' surprisal when a speaker switches to a different language. Results showed no evidence that infants spontaneously associate a person with a language. This contrasts with common but outdated advice to caregivers to choose a single language when speaking to their child, and is consistent with research showing that bilingual infants learn languages from a variety of family language strategies.

Manuscript 3 examined how bilingual infants mentally represent the sounds in familiar words. Bilingual and monolingual infants were tested to examine whether bilingual infants are more or less sensitive to mispronunciations than monolinguals. I tested 24- to 26-month-olds in a looking-while-listening task. I found a robust mispronunciation effect in bilingual and monolingual toddlers, indicating that bilinguals' encoding of sounds in familiar words is phonetically detailed in a similar way as monolinguals', despite bilinguals having to navigate a more complex phonetic environment.

In these three manuscripts, I found that bilingual language acquisition is similar to monolingual development in many aspects, and rigorous testing of assumptions about bilingual language acquisition is needed to learn about the mechanisms bilingual infants use to acquire language.

Acknowledgements

Many people have supported me on this journey, and I want to extend my deepest and enduring thanks to this community. Krista, I will forever be grateful for your mentorship. Thank you for providing me with the space to grow, as a researcher and as a person, and for cheering me on at every point of my academic journey. You have created a wonderful lab environment, and I strive to be the mentor you are to others around me. I would also like to thank my committee members, Diane Poulin-Dubois and Kristen Dunfield, for sharing their enthusiasm for developmental research with me, and providing constructive feedback at many stages of this project. Many thanks also goes to the members of the Cognitive and Cognitive Development and Bilingualism Interest Groups in the Department of Concordia, who have created a nurturing academic exchange. I also want to thank my mentors from previous stepping stones in my academic career, including Janet Dean Fodor, Debra Titone, Maria Vuori, and Karsten Stegmann.

My heartfelt thank you also goes to the families who generously volunteered their time to participate in research, and graciously shared a glimpse of how they raise their children. In addition, I want to thank the staff in the Concordia Psychology department, who always were helpful in navigating graduate school.

A huge part of my graduate school experience were the members of the Concordia Infant Research lab. I want to thank the project students I've had the pleasure of working with, Maria Tamayo, Meghan Mastroberardino and Alexa Fogel, thank you for bringing your enthusiasm into our research projects. A big thank you goes to past and present lab members, in no particular order, Lena Kremin, Laia Fibla Reixachs, Melanie López Pérez, Nicolás Alessandroni, Erin Quirk, Ana Maria Gonzalez Barrero, Alexandra Laoun-Rubenstein and Daphnée Dubé. A special thank you goes to Hilary Killam, for being a great friend and a stellar coworker, with an eye for detail and spot on feedback at every stage. Thank you also to Elizabeth Morin-Lessard, for paving the way for those who came after you and for sharing your passion for your work. Rachel Tsui, thank you for always being up for ice cream and adventures. Charlotte Moore, thank you for your sharing your enthusiasm for dogs, biking, and research. Your advice and entertainment recommendations kept me sane during the long hours of thesis writing. Andrea Sander Montant, I will miss our lunch dates so much, thank you for being always there for me! Melanie Brouillard, our pandemic walks kept me same, thank you for being an awesome friend.

Thank you also to my writing communities who have made the process of writing this thesis more enjoyable, and when it wasn't, were always there to listen. Countless thanks goes to my virtual office mates in the last months of thesis writing, Joscelin Rocha-Hidalgo, Carla Macias, and Kareem Willis, I'm so grateful I got to go through this experience with you! A huge thank you also goes to the members of 100daysOfWriting, for making writing a communal activity, and always cheering each other on.

There are numerous other people who have been part of this journey, including Karim Bouayad-Gervais, Geneviève Bouchard, Erin Yong Ping, Sonia Krol, Kayleigh-Ann Clegg, Colin Stewart, thank you for all the laughs we shared. Thank you to Jonas Nitschke and Nida Ali, for their support during all parts of my academic journey, and the best cabin weekends. Thank you to Julie Arseneau and Mitch Kibsey, for being miracle pandemic friends always up for an adventure. A heartfelt thank you goes to Jessica Hamel-Akré, for so many helpful chats about life, academia, and sewing. Thank you to my brother Chris, for some of my favourite memories I made in Canada on our road trips.

Thank you to my mom, who instilled in me a love of languages. Thank you to my father, who taught me that even when a problem seems hard, there's usually a solution out there.

Finally, thank you to Sebastian Urchs, for being by my side supporting me, and for growing together.

I dedicate this thesis to my grandmothers, who have shown me that a woman's strength comes in many forms.

Contribution of Authors

This dissertation consists of three manuscripts containing five studies, as follows:

Manuscript 1 (CHAPTER 2; comprising two studies)

Schott, E., Mastroberardino, M., Fourakis, E., Lew-Williams, C., & Byers-Heinlein, K. (2021). Fine-tuning language discrimination: Bilingual and monolingual infants' detection of language switching. *Infancy*, *26*(6), 1037-1056. https://doi.org/10.1111/infa.12429

ES: Conceptualization; Investigation; Data Curation; Formal Analysis; Visualization; Writing – Original Draft Preparation; Writing – Review & Editing
MM: Conceptualization; Data Curation; Visualization; Writing – Original Draft Preparation
EF: Investigation; Data Curation
CLW: Conceptualization; Funding Acquisition; Supervision; Writing – Review & Editing
KBH: Conceptualization; Funding Acquisition; Supervision; Writing – Review & Editing

Manuscript 2 (CHAPTER 3; comprising two studies)

Schott, E., Tamayo, M. P., & Byers-Heinlein, K. (under review). *Keeping track of language: Can monolingual and bilingual infants associate a speaker with the language they speak?* Manuscript submitted for publication.

ES: Conceptualization; Data Curation; Formal Analysis; Visualization; Writing – Original Draft Preparation; Writing – Review & Editing
MPT: Conceptualization; Data Curation; Visualization; Writing – Original Draft Preparation; Writing – Review & Editing
KBH: Conceptualization; Funding Acquisition; Supervision; Writing – Review & Editing

Manuscript 3 (CHAPTER 4; comprising one study)

Schott, E., & Byers-Heinlein, K. (in preparation). Bilingual and monolingual toddlers are sensitive to mispronunciations for familiar cognate and non-cognate words.

ES: Conceptualization; Data Curation; Formal Analysis; Visualization; Writing – Original Draft Preparation; Writing – Review & Editing
KBH: Conceptualization; Funding Acquisition; Supervision; Writing – Review & Editing

All authors reviewed the final manuscripts and approved of the contents.

Table of Contents

List of Tables

List of Figures

1 Introduction

About 1 in 6 children living in Canada grow up hearing two languages (Schott et al., 2022). This number goes up to 1 in 4 children for those living in Canadian cities such as Montréal and Vancouver. Bilingual children learn to understand and use two languages in the span of a few years, all without explicit awareness that they are hearing multiple languages. Bilingual children face particular challenges stemming from their dual exposure that are not present for monolingual infants. This dissertation focuses on three questions relevant to bilingual children's early language development. These three questions will be addressed in a series of empirical studies testing young children, organized into three chapters. In each study, I tested bilingual and monolingual children, to investigate whether bilinguals' lifetime language exposure affected their performance, or whether regardless of language exposure, children of a particular age group perform similarly on a task. In Chapter 2, I investigated whether bilingual and monolingual infants can detect language switching at short timescales, for example inside a sentence or list of words. In Chapter 3, I study whether bilingual and monolingual infants notice when a person switches to a different language, which could be a useful strategy to keep track of their languages in a changing language environment. In Chapter 4, I tested toddlers' word recognition to test whether the properties of the language input affect how words are recognized and encoded. In the remainder of this chapter, I will provide an overview of bilingual language development, focussing on language discrimination, indexical information in the child's language input, and word recognition, and then provide a detailed outline of the dissertation's objectives.

1.1 Early language discrimination

The question of when infants can discriminate between different languages has received much attention. Investigating this question can tell us about what infants learn about the languages in their environment (for a review, see Gasparini et al., 2021). The speech that infants encounter varies on many characteristics such as the phonology and prosody. Much of the language discrimination literature asks which of these aspects infants pay attention to when discriminating languages. The earliest research done on this was mostly focussed on testing monolingual infants using long passages of speech that alternate between languages. One example is a study of French newborns (Nazzi et al., 1998). When tested on two different-sounding languages, such as English and Japanese, newborns were able to discriminate them. When tested on two similar-sounding languages, such as English and Dutch, newborns did not show discrimination.

A recurrent finding in the early language discrimination literature was that some language pairs are discriminated earlier than others. For example, French newborns were unable to discriminate English and Dutch, but were able to discriminate English and Japanese (Nazzi et al., 1998). These results prompted theories about what properties make languages easier or harder to discriminate. Languages differ in many aspects, including the sounds present, prosody and language rhythm. Initial theories focussed on the idea that languages can be categorized into different rhythmic classes. For example, English and German are stress-timed, whereas many Romance languages are syllable-timed (Grabe & Low, 2002). Languages that are in the same rhythmic class were thought to be harder to discriminate, and languages from two different rhythmic classes were thought to be easier to discriminate. However, classifying languages into rhythmic classes is not a straightforward endeavour (Grabe & Low, 2002; Ramus, 2002). A recent review found that a better predictor of whether infants will discriminate two languages is cross-language differences in variability of vowels and consonants (Gasparini et al., 2021). To summarize, infants can discriminate some but not all languages from birth, and language pairs differ in how easy or hard it is to discriminate them. Durational variability in language rhythm plays an important role for understanding which pairs are discriminated more easily.

The ability to discriminate languages is likely also important for bilingual infants. Bilingual infants, by definition, hear two languages in their everyday life. If they are able to discriminate their languages, they might use their language discrimination abilities to 'tag and sort' the language input they receive into separate representations, one for each language (Byers-Heinlein, 2014; Nazzi et al., 1998; Sundara & Scutellaro, 2011). Despite the question of language discrimination being very relevant for bilingual infants, much of the initial research on language discrimination tested monolingual infants. There have, however, been some studies on bilingual infants' ability to discriminate languages. Newborns who have been exposed to English and Tagalog in their mothers' womb can discriminate between English and Tagalog (Byers-Heinlein et al., 2010). Several studies tested slightly older bilingual infants' ability to discriminate more similar-sounding languages, such as Spanish and Catalan or Spanish and

Basque and found that 3-5 month-olds were able to discriminate their languages (Bosch & Sebastian Galles, 2001; Molnar et al., 2014; Nacar Garcia et al., 2018). Based on the available literature on language discrimination in the first year of life, we would expect that at least by 5 months, infants are able to discriminate most language pairs.

Yet, bilingual language environments are much more nuanced than the language discrimination tasks used in these discrimination studies. In the studies discussed above, infants hear multiple sentences in one language, and then multiple sentences in the other language. In bilingual environments, some parents switch one or more times within a single sentence (Kremin et al., in press). If infants keep track of what language is spoken, these rapid language switches present a potential problem, as shorter intervals of speech contain fewer cues to which language is being spoken. This is particularly important as theorists have emphasized the role of durational variability in longer segments of speech (Gasparini et al., 2021; Ramus, 2002). As discussed above, being able to detect whether even a single word comes from a different language would be useful for infants to sort the words they learn into language 'categories' (Byers-Heinlein, 2014). In line with this suggestion, studies have found that at least by 20 months, infants are able to detect single-word language switches in a comprehension task (Byers-Heinlein et al., 2017; Potter et al., 2018, 2019). In this task, infants see an image of a dog and hear "look at the *chien* [fr. dog]" and their looking towards the target compared to the distractor is tested. Thus, infants in this task probably rely strongly on their lexical knowledge. However, infants in their first year of life cannot only rely on lexical knowledge to identify the language being spoken, as they know few words at this age. Thus, it is unclear whether younger infants would detect single-word language switching. In Chapter 2, I will address this gap in the literature by testing whether 8- to 12-month-old infants can detect single-word language switches.

1.2 Person-language associations in bilingual development

Bilingual families vary on many dimensions, and the ways in which bilingual infants are exposed to their languages also varies (e.g., De Houwer, 2007; Kremin et al., in press). A popular approach for exposing bilingual infants to their languages is the "one-parent-one-language" approach. The idea of this approach is that the consistent use of one language by each parent would help bilingual children separate their two languages (Barron-Hauwaert, 2004). While relatively few children hear a strict version of the "one-parent-one-language" approach (De Houwer, 2007), it is common for parents to use predominantly one language with their child in a way that is relatively consistent over time (Orena et al., 2020). In this case, bilingual children could benefit from their ability to discriminate between people (Johnson et al., 2011) to keep track of what language is spoken. In the following section, I will review some of the findings on monolingual infants' abilities to use language cues to reason about person characteristics and vice versa.

Infants learn about the world by observing those around them, and there is some evidence that infants use information about who speaks what language to guide their interactions with the world. For example, when monolingual 10-month-olds were given the chance to take a toy from a person who spoke their native language or an unfamiliar language, they took the toy from the speaker of their native language (Kinzler et al., 2007). This also extends to infants' tendency to imitate others, where 14-month-olds were more likely to imitate speakers of their own native language compared to speakers of an unfamiliar language (Buttelmann et al., 2013; but see Howard et al., 2014). Infants can rely on shared language use between two people to make inferences about their social affiliation (Liberman et al., 2017). When observing two actors speaking different languages, monolingual 9-month-olds looked longer when the speakers interacted positively compared to when they interacted negatively. The same pattern was not observed when the actors spoke the same language. These findings suggest that monolingual infants can use information about what languages a person speaks to guide their own actions and to reason about others' social affiliation. This has not been tested in bilingual infants.

We can also gain insights from studies testing another aspect of visual cues that infants pay attention to when looking at people, in particular visual cues to a person's race. Previous research has shown that infants' reasoning about language intersects with their perception of a person's race. For example, a study of monolingual 10-month-olds showed that infants looked more to the face of an Asian woman compared to a white woman when hearing Cantonese, but showed no matching between language and speaker characteristics when hearing English (May et al., 2019). These infants were growing up in Vancouver where both languages are spoken in the community, and this suggests that infants had learned an association between Asian faces and hearing Cantonese. Further evidence that infants are sensitive to a person's race in the context of language comes from a study where monolingual white 16-month-olds growing up in Waterloo, Ontario (a community with fewer visible minorities than Vancouver) heard familiar and unfamiliar accents that were paired with a picture of a woman of the same race or a different race as the infant (Weatherhead & White, 2018). The authors found that infants who saw the same-race speaker expected to hear a familiar accent, and looked less at the target when hearing an unfamiliar accent, but for the other-race speaker no such difference between accents was found. This paradigm was also used to test bilingual and monolingual 25-month-olds growing up in Singapore, a multi-racial society (Singh et al., 2020). Monolingual infants showed different looking patterns when they saw the other-race speaker compared to the same-race speaker like in the previous study. Interestingly, bilingual infants showed the same looking patterns for both speakers. Taken together, these findings show that monolingual infants from a variety of communities do take visual cues to the race of a person into account when processing language. Much less research exists about bilingual infants, but the one study that has tested this population (Singh et al., 2020) suggests that bilinguals may not use visual cues to race in language processing.

Overall, the studies reviewed in this section indicate that monolingual infants can learn about a person's language use and use that information to change their own behaviours or reason about other's actions. Bringing this back to the context of bilingual language learning in infants, this ability could be beneficial to infants faced with a rapidly changing language input, for example due to rapid language switching, where they could use a person's identity to find person-specific patterns of language use. This would be useful for infants exposed to languages in a person-specific manner, for example using the one-parent-one-language approach. However, the ability of infants to associate a person and a language in a spontaneous manner has not been tested empirically. In Chapter 3, I will investigate this issue in 5- to 18-month-old bilingual and monolingual infants.

1.3 Early Bilingual Word Comprehension: Do Properties of Words Matter?

A particularity about bilingual word learning is that bilinguals regularly learn two words for a concept, such as "apple" and "*pomme"* (fr. apple). These words that share meaning are called translation equivalents. Furthermore, some translation equivalents also share sounds, such as "chocolate" and "*chocolat"* (fr. chocolate). These words that share meaning and also sounds and/or orthography are called cognates. In the adult bilingualism literature, cognates have often been used to probe whether both languages are activated during reading (Assche et al., 2012). A common finding is that bilingual adults process cognates faster than non-cognates in reading tasks. One explanation for this is that when seeing a cognate, the overlap between word forms causes both word forms to become activated, which speeds word recognition for cognates relative to non-cognates (Dijkstra et al., 2010). While in young children orthographic effects of cognates are irrelevant as they have not yet learned to read, the cross-language overlap in sounds present in cognates may still affect how bilingual children learn and comprehend words.

There is some evidence that cognates and non-cognates are processed differently in children. In a review paper on studies that included children from 3 to 8 years, most but not all studies found differences between cognate and non-cognate words in a range of tasks including picture naming, and translation (Squires et al., 2020). Other studies suggest that cognates are acquired faster or more easily compared to non-cognate words (Mitchell et al., 2022; Schelletter, 2002). However, the effect of cognates on language processing may be modulated by the characteristics of the bilingual experience, such as how much exposure bilingual children have to each language and when that exposure starts. One study tested whether cognates are recognized faster compared to non-cognates in German-English bilingual 1.5- to 4.5 year-olds (Von Holzen et al., 2018), and found that this was only the case for English (the children's second language), but not for German (the children's first language). In a study using accuracy in pointing to pictures of cognate and non-cognate words (Pérez et al., 2010), children dominant in Spanish (and tested in English) were more accurate at pointing to cognates compared to non-cognate words. The opposite pattern was found for English-dominant children, who were less accurate at pointing to non-cognates compared to cognates. No difference was found for children who had balanced exposure to two languages. Further research is needed to explore the interaction of different experiences of bilingualism and cognate status, however we have tentative evidence that cognates play a larger role in bilingual children dominant in one language. In Chapter 4, I will test bilingual toddlers' ability to recognize cognate and non-cognate words, to test whether cross-language similarity affects their word recognition.

1.4 Early (Bilingual) Word Comprehension: How Detailed Are Word Representations?

An important aspect of language acquisition is for children to learn to understand the meanings of the words that are spoken around them. In learning to recognize words, children must learn that some of the variability inherent in language does not change meaning, such as when a child hears the word "ball" pronounced slightly differently by a younger or older speaker, or in different intonations. At the same time, children must learn that some types of sound variability do change meaning, for example when hearing "bowl" and "ball", where the vowel just changes slightly and the meaning is different. This is a potentially difficult feat for infants learning about their languages' sounds and words at the same time, particularly for bilinguals who are acquiring two languages.

To test how sensitive infants' word recognition is to small changes in pronunciation, a mispronunciation paradigm is often used (Swingley, 2005, 2009; Von Holzen & Bergmann, 2021). Infants see two objects on a screen, for example a table and a banana, and hear one of them labelled, either correctly ("Look, a banana!") or incorrectly ("Look, a boonana!"). The proportion of time they spend looking at the target object (here, the banana) compared to the distractor object (here, the table) indexes how well and efficiently they process and understand the word. There are three possible outcomes in this task when infants are tested on correctly pronounced and mispronounced words. First, infants could look equally to the target when hearing "banana" and "boonana". This would suggest that they did not perceive the difference between the two labels, or that their word representation was not detailed enough to notice the mismatch when hearing the mispronunciation. Second, infants could look at the target both when hearing "banana" and "boonana", but do so significantly less when hearing "boonana" compared to "banana". In this case, we would infer that infants' perceptions of the sounds in the label and their word representation were detailed enough to notice that "boonana" is not quite the right label for a banana, but flexible enough to infer that the speaker probably meant to refer to the banana. A third possibility would be that infants would only look at the banana when hearing the correctly pronounced label "banana", and not look above chance at the target object when hearing "boonana". This would indicate a detailed word representation, accompanied by a rigid interpretation of the sounds in a given word that fails to recognize a word when it is slightly mispronounced. Thus, the mispronunciation paradigm can be used to test how children cope with variability in word recognition.

When tested in a mispronunciation paradigm, monolingual infants show early sensitivity to mispronunciations, starting at around 11 months (Bergelson & Swingley, 2018), and this ability continues through early language development (for a review, see Von Holzen & Bergmann, 2021). Typically, the results from monolingual children match the second outcome discussed above most closely, where they are above chance in both correctly pronounced and mispronounced trials, but look less at the target in mispronounced trials compared to correctly pronounced trials (Von Holzen & Bergmann, 2021). This suggests that monolingual infants have a detailed phonological representation, and often can strike a balance between leniency to non-meaningful variability in word comprehension and attention to meaningful variability, such as mispronunciations.

Like monolinguals, bilingual infants are also navigating this balance between leniency and attention to detail, yet there have been fewer studies exploring mispronunciation sensitivity in bilinguals than in monolinguals. Bilingual infants are potentially exposed to more variable input, as they are more likely to hear input from other bilingual people who show increased variability in pronunciations, and this could change bilingual infants' speech perception (Byers-Heinlein & Fennell, 2014). The few studies that tested bilinguals' sensitivity to mispronunciation found somewhat inconclusive results, mostly showing that bilingual infants are sensitive to mispronunciations (Singh et al., 2020; Tamási et al., 2016; Wewalaarachchi et al., 2017). However, one study found that bilinguals sometimes are not sensitive to certain vowel mispronunciations, in particular ones that are phonemic in only one of their languages (Ramon-Casas et al., 2009). In a set of related studies testing infants' ability to use subtle phonetic contrasts in a word learning study, bilingual infants learned the nonsense words "bih" and "dih" as labels for novel objects later than monolinguals (Fennell et al., 2007). In this study labels were produced by a monolingual speaker, however in follow up studies that used stimuli produced by bilingual speakers for the bilingual infants and monolingual speakers for monolingual infants, both groups succeeded at the same ages (Fennell & Byers-Heinlein, 2014; Mattock et al., 2010).

The specific combination of languages bilingual infants are learning also appears to impact their ability to use subtle phonetic contrasts. In one study, bilingual infants learning two languages where phonemic boundaries were similar succeeded in learning novel words, whereas bilingual infants learning languages where phonemic boundaries were different failed to learn the novel words (Havy et al., 2016). These findings highlight that bilingual infants' ability to use subtle phonetic differences depends on the task at hand, showing the need for further research to uncover when bilingual infants are able to use these types of information. Chapter 4, in addition to assessing cognate effects, will test bilingual infants' ability to use subtle phonetic differences in a mispronunciation task.

1.5 Dissertation Research Objectives

Bilingual language learning is a multi-facetted process that is important to study to understand how bilingual and monolingual children learn languages. In investigating four related research questions on bilingual language acquisition as described in the previous sections, this dissertation addresses two main objectives to advance our understanding of early language learning.

The first objective is to investigate infants' ability to discriminate languages in situations that are typical in bilingual infants' early life. Bilingual infants are often exposed to rapid language switching, including single-word language switches and people switching between languages, which may pose a challenge to infants trying to acquire the words for each language separately. In Chapter 2, I tested whether 8-12-month-olds can discriminate between single-language and switched-language utterances. Furthermore, each caregiver of bilingual infants may speak one or more languages systematically, and thus to keep track of which language is spoken, it may be helpful for infants to be able to make person-language associations. In Chapter 3, I familiarised 5-18-month-olds to speakers who speak one language consistently, and then tested whether infants are surprised when the speaker switches to another language. By understanding when and if infants discriminate languages, we can learn what infants pay attention to in their language environment and which strategies they use to acquire language.

The second objective is to investigate some factors that may affect how children represent and understand words. Being able to communicate with the world around them is an essential milestone of language learning, and therefore I studied toddlers' spoken word comprehension. In Chapter 4, I tested 24- to 31-month-olds on their recognition of familiar cognate and non-cognate words. Some of the words are mispronounced, to probe how phonologically detailed toddlers' word representations are.

The Discussion section (Chapter 5) explores how the contributions of each paper intersect and inform each other. More generally I discuss the importance of null-results in developmental research, and the nuances of comparing bilingual and monolingual children's language skills. Overall, this thesis contributes to our understanding of how bilingual and monolingual children learn language.

2 Fine-tuning language discrimination: Bilingual and monolingual infants' detection of language switching

2.1 Introduction

All bilingual environments involve periodic switching between languages either within or across speakers, and an ability to detect these switches is foundational to bilingual infants' successful language acquisition and later language use. Previous research has tested infants' detection of language switches at longer time scales, such as following narrative-like passages (e.g., Bosch & Sebastian Galles, 2001; Nazzi et al., 1998). However, little is known about infants' ability to detect language switches at shorter time scales. Given that language learning in the first year of life is largely focused on sequences of sounds and words, how do infants process switches at the level of individual words? The ability to detect a transition between languages in a sentence such as "Look at the *chien* [fr. dog]" would presumably support bilingual infants' emerging representations of words in two languages (Byers-Heinlein, 2014; Curtin et al., 2011). The present studies tested whether bilingual and monolingual 8- to 12-month-old infants, who are just beginning to learn words, can detect single-word language switches.

Language switches are frequent in many bilingual infants' language environments. While the amount of language switching varies across families and communities, most bilingual parents switch languages in interactions with their children, with some children hearing language switches in as many as 2/3 of utterances (Bail et al., 2015; Byers-Heinlein, 2013; Kremin et al., in press). Language switching can take several forms. First, parents may alternate their languages at the point of a sentence boundary (e.g., "*Regarde ici* [fr. look here]! Do you see the dog?"). Second, they may switch languages within a sentence by borrowing a single word from another language ("Do you want your *toutou* [fr. stuffed animal]?"). Finally, they can switch across isolated words, such as when teaching new words (e.g., "Look! Dog! Chien [fr. dog]!" Byers-Heinlein, 2013). Some of these types of language switches may be more difficult for infants to detect than others.

Most research on infant language discrimination has focused on cross-sentence switching in long passages of speech. In one study, newborns born to bilingual mothers, who were exposed to both English and Tagalog during pregnancy, were tested in a habituation paradigm (Byers-Heinlein et al., 2010). After hearing sentences in one of their maternal languages until they lost interest, newborns showed renewed interest only when the stimuli switched to the other language, suggesting that they could detect the language change. Thus, following prolonged exposure to a language in a long passage, even newborn bilinguals are able to detect language switches for some language pairs. More evidence for infants' ability to discriminate languages across longer time scales comes from studies of Spanish–Catalan bilingual 4-month-olds (Bosch & Sebastian Galles, 2001) and Spanish–Basque 3.5-month-old bilinguals, as well as studies including monolingual infants, from birth to 5 months of age (Bahrick & Pickens, 1988; Nazzi et al., 1998, 2000). In these studies, language rhythm, an aspect of prosody that is largely related to variation in consonant and vowel duration, appears to play a role in detecting language switches (Gasparini et al., 2021). Rhythmic information is richer in multisyllabic utterances (Ramus et al., 1999), thus, these studies leave open the question of whether infants can detect language switches over shorter time scales where rhythmic cues are less available.

Some recent evidence suggests that at least older infants might be able to detect language switches across shorter time scales, for example when switches involve a single word (e.g., "Do you see the *chien* [fr. dog] over there?"). In an eye-tracking study (Byers-Heinlein et al., 2017), 20-month-olds looked less at a target image when the target word was language-switched ("Look at the *chien*") compared to when the target word was of the same language ("Look at the dog"). Similar results were found for English–Spanish bilingual 18- to 30-month-olds (Morini & Newman, 2019; Potter et al., 2018, 2019). Additional evidence for detection of language-switched words comes from a study of English–Welsh bilingual 2- to 3-year-olds' event-related potentials (Kuipers & Thierry, 2012). Together, these studies suggest that by their second year of life, bilingual infants have some ability to detect a language switch at the level of individual words.

Even during the first year of life, infants possess prerequisite abilities that could help them to detect single-word language switches. While direct evidence for this is sparse, several studies indicate that infants' budding knowledge of sounds and words could enable them to detect such variation. First of all, monolingual infants learn about the sound patterns in their native language prior to the onset of word production (Werker, 2018), and 6- to 9-month-olds show precursors of word comprehension for frequent words by looking at a labelled referent (Bergelson & Swingley, 2012, 2015; Kartushina & Mayor, 2019). While such studies have yet to be done with bilingual infants, ample evidence suggests that bilingual infants make rapid gains in learning about the sounds (i.e., phonetic inventory) and allowable combinations of sounds (phonotactics) of their two languages. Within the first year, both monolinguals (Kuhl et al., 2006; Werker & Tees, 1984) and bilinguals (Albareda-Castellot et al., 2011; Burns et al., 2007; Sundara & Polka, 2008) become perceptually specialized to the sounds of their native language, and further, bilingual infants are sensitive to patterns of sounds that are typical in their dominant language (Sebastian Galles & Bosch, 2002). Moreover, bilingual infants start to recognize the sound patterns of frequently-heard familiar words by 11 months (Vihman et al., 2007), though as mentioned above more studies are needed on bilinguals' early word comprehension. Thus, it is possible that even young bilinguals could exploit their developing knowledge of sounds and sound patterns to detect single-word language switches, at least for certain highly frequent and familiar words.

Here, in two studies, we investigated whether infants are able to detect single-word language switches. We tested 8- to 12-month-old infants, because at this age they have already acquired knowledge about the sounds and sound patterns in their native language(s) and have started to learn about the words in their language(s) as well. Infants listened to single-language and switched-language speech, and their listening times (operationalized as looking times) were measured using the head-turn preference procedure. Words were presented in a word list in Study 1 or following a naturally produced sentence frame in Study 2, which allowed us to test whether the local context of a language switch matters for infants' detection of single-word language switches. On single-language trials, infants heard only one language, e.g., "dog… milk… dog… milk…" in Study 1 or "Do you like the dog? I want the milk!" in Study 2. On switched-language trials, infants heard words from two languages, e.g., "dog… *lait* [fr. milk]… dog… *lait*…" (Study 1) or "Do you like the *chien* [fr. dog]? I want the *lait*!" (Study 2). The same infants participated in both studies during the same lab visit.

We predicted that infants would succeed in detecting language switches both in word lists (Study 1) and in sentences (Study 2), indexed by significantly different looking times to single-language and switched-language trials, although we did not predict a direction of the difference given the challenge of doing so for a new paradigm (Aslin, 2007). It was also possible that infants would only succeed at detecting language switches when embedded in a sentence. Multi-word utterances are more common than single-word utterances in input to children (Brent & Siskind, 2001), which may facilitate infants' using their day-to-day language processing abilities more fully for sentence-embedded (Study 2) compared to single-word (Study 1) language switches. Furthermore, most previous studies of early language discrimination have used stimuli that contained long passages of speech, and from this research we know that rhythmic cues are important to language discrimination (e.g., Ramus, 2002). In a similar vein, bilingual toddlers were better at recognizing familiar words when they were embedded in a sentence than when they were heard in isolation (Morini & Newman, 2019). Hearing *"Do you like the…"* provides infants with more exposure to sequences of sounds and words in the language being spoken prior to a language switch.

We tested both bilingual and monolingual infants. While we expected all infants to detect the switches, we expected that bilinguals might show a larger effect (i.e., a larger looking time difference) than monolinguals. This would indicate that everyday exposure to two languages enhances this ability. Another possible outcome would be that bilingual and monolingual infants show equivalent detection of language switches, suggesting that this ability is not related to the experience of hearing both languages. Testing both bilingual and monolingual infants allowed us to investigate how language exposure interacts with infants' ability to detect language switches.

2.2 Study 1: Language Switches in Word Lists

2.2.1 Method

The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. This study was approved by the Concordia University Human Research Ethics Board (certificate #10000439) and the Princeton University Behavioural Research Ethics Board. All parents provided informed written consent prior to their infants' participation in the study (see Appendix A). Materials and methods are available at https://osf.io/9dtwn/.

2.2.1.1 Participants

Participants were 21 English–French or English–Spanish bilingual infants, as well as 20 monolingual infants. Instead of a power analysis, we conducted a sensitivity analysis, which is useful when there are constraints on time and participant recruitment (Lakens et al., 2018; e.g., for hard-to-recruit populations like bilingual infants, Schott et al., 2019). Using a sensitivity analysis in G*Power (Faul et al., 2007), we found that a mixed ANOVA with a sample size of 20 infants per group would yield 80% power to detect an effect of Cohen's d = .45. This value was used as our minimum effect size of interest for equivalence tests (presented in the Results section), and is comparable to the average effect size found in a meta-analysis of familiar word recognition studies using similar stimuli and testing mostly monolingual samples of the same age (Cohen's d = 0.54, retrieved from *MetaLab Project*, 2020; see also Carbajal et al., 2021). Our original research plan was to test bilingual infants only, but we were able to further explore the effect of language background when an opportunity arose to test monolingual children in Montréal, Canada. Bilingual children were recruited in both Montréal, Canada (English–French bilinguals, $n = 14$) and in New Jersey, USA (English–Spanish bilinguals $n = 7$). The language environment in these two communities is quite different: In Montréal, both English and French are widely spoken in the community, while in New Jersey, English is the majority language. The original research plan involved testing a sufficient number of bilingual infants in both locations to enable a comparison, but due to recruitment difficulties, only a limited number of infants in New Jersey met the language inclusion criteria articulated below and could be tested. Thus, the data from both locations were combined. Further information on age and gender is displayed in Table 2.1.

Infants' language exposure was assessed using the Multilingual Approach to Parent Language Estimates (Byers-Heinlein et al., 2020) which assesses infants' month-by-month exposure to languages from birth to the test date (for the form used, see Appendix B). Infants were considered monolingual if they were exposed to 90% or more of English or French, and bilingual if they were exposed to 25-75% of English and 25-75% of French or Spanish. The language that infants were exposed to the most was considered their dominant language, and the percentage of infants dominant in English is reported in Table 2.1. On average, bilingual infants tested in Montréal heard their dominant language 63% (range: 42–75%) and their non-dominant language 36% (range: 26–50%) of the time. Bilinguals tested in New Jersey heard their dominant language 57% (range: 51–73%) and their non-dominant language 43% (range: 27–49%) of their time. Monolinguals tested in Montréal heard their dominant (native) language on average 98% (range 91–100%) of the time. Two bilinguals had exposure to a third language (13% and 20% exposure, respectively). The Language Mixing Scale (Byers-Heinlein, 2013) was used to assess exposure to parental language mixing (see Appendix C), and bilingual infants heard more mixing than monolingual infants (bilingual infants: 13.05, $SD = 7.39$, range = [0–25]; monolingual infants: 5.55, $SD = 5.01$, range = [0–16]; $t[37] = 3.73$, $p = 0.001$, Cohen's $d = 1.16$). All participants were reported to have normal vision and hearing. Infants and their families were recruited through government-supplied birth lists, as well as in daycare centers, playgroups, and other child-focused community activities. Infants in New Jersey were tested between August 2016 and November 2017 and infants in Montréal were tested between February 2017 and November 2018.

Table 2.1 Information about Infant Participants in Studies 1 and 2, by Language Group

		Age				
Language Background	N	Min	Mean	Max	% Male	% English dom.
			Study 1			
bilingual	21	7m 29d	9m 19d	12m 16d	33	48
monolingual	20	7m 29d	8m 24d	10m 2d	50	30
			Study 2			
bilingual	20	7m 29d	9m 6d	11m 17d	40	55
monolingual	17	7m 29d	8m 25d	10m 2d	47	29

Note. All infants were tested in both studies in the same lab visit, but in a few cases infants only successfully completed one study. For monolinguals, the column % English dominant denotes the percent of children who are English monolinguals. As monolingual children were only tested in Montréal, children who are not English-dominant are French-dominant.

To achieve the final sample, 68 infants and their families participated in the study. Two participants were tested during a pilot phase to verify that the study procedure was feasible, and were not included in data analysis. Of the remaining infants, one was born prematurely (< 37 weeks) and thus did not meet our health criteria, and 16 did not meet the pre-specified language criteria. Four infants were excluded for technical difficulties or experimenter error. Additionally, four infants were excluded for contributing fewer than 8 trials with at least 2.5 s of looking time. Trials shorter than 2.5 s were excluded because infants needed to listen at least that long in order to encounter a switched-language word. Some infants completed fewer than 16 trials because they were fussing and their parent ended the study early, but in these cases, infants were not automatically excluded from analysis and all usable trials were analyzed. There were no other exclusion criteria.

2.2.1.2 Stimuli

In the head-turn preference procedure, visual stimuli were presented on both centre and side screens to attract infants' attention. The visual stimuli differed between the Montréal and New Jersey test locations due to different lab conventions. Infants tested in Montréal saw an animation of a spinning rainbow-coloured citrus, infants tested in New Jersey saw a video recording of an orange flashing light. This was not expected to affect the pattern of results, as the visual stimulus only served to attract infants' attention, and was constant across auditory stimuli.

Speech stimuli for infants tested in Montréal were recorded by a native English–French bilingual, and for infants tested in New Jersey by a native English–Spanish bilingual. All words were non-cognates, and were chosen because they are, on average, acquired early in language development. Word pairs presented together in a trial were always of the same grammatical gender, and were thematically dissimilar (Willits et al., 2013). To the extent possible, we avoided overlapping word onsets or codas, and we matched word pairs on number of phonemes, word frequency, and stress patterns (which was not possible for English–French disyllabic words due to differences in typical stress patterns across languages). Due to these constraints, the word pairs in the English–French and the English–Spanish versions of the study are partially but not fully overlapping. For the English–French version of the study, word pairs were dog - milk, kitty - book, mouth - door, cookie - foot and their French translations; for the English–Spanish version, word pairs were doggy - balloon, kitty - foot, mouth - milk, cookie - door and their Spanish translations. Examples of the stimuli used in the study can be seen in Table 2.2, and all original stimuli can be seen in Table E1 in Appendix E, and downloaded at https://osf.io/9dtwn/.

Each word was recorded separately, in a friendly, infant-directed manner, and then combined to form single-language (e.g., dog…milk) and switched-language trials (e.g., dog…*lait*). Speakers were asked to produce a consistent, hill-shaped prosodic contour across items (Nencheva et al., 2021), and we selected tokens that sounded similar in their prosody across languages. Individual words were presented in alternation with a 500 ms pause between each word. Trials lasted 20.4–22.6 s depending on the length of the audio file. All looking times were capped at the shortest trial length to avoid introducing a difference in looking time based on the length of the audio file.

Table 2.2 Examples of Single- and Switched-Language Trials in Studies 1 and 2

Language Pair	Version	Single-Language	Switched-Language
		Study 1	
English–French	English	Dog… Milk…	Dog… *Lait*…
	French	Chien… Lait…	*Chien*… Milk…
English–Spanish	English	Doggy… Balloon…	Doggy… *Globo*…
	Spanish	Perro… Globo…	*Perro*… Balloon…
		Study 2	
English–French	English	Do you like the milk? I want the dog!	Do you like the *lait*? I want the *chien*!
	French	Aimes-tu le lait? Je veux le chien!	Aimes-tu le milk? Je veux le dog!
English–Spanish	English	Do you like the balloon? I want the doggy!	Do you like the *globo*? I want the *perro*!
	Spanish	¿Te gusta el globo? Quiero el perro.	¿Te gusta el balloon? Quiero el doggy.

Note. Each participant heard only one version (e.g., an English–French bilingual was randomly assigned to hear either the English or the French version of the study). Non-English words are in italics.

2.2.1.3 Procedure

Infants sat on their caregiver's lap in a sound-attenuated room with three monitors, one centered, one to the left, and one to the right of the infant. The caregiver listened to music through a set of headphones to avoid influencing the infant's reactions. The experimenter controlled the study via custom Matlab software (Olson, 2017). The experimenter was in a different room for infants tested in Montréal or in the same room but listening to masking music through headphones for infants tested in New Jersey. In both cases, the experimenter was unaware of the experimental condition on each trial. At the start of each trial, the visual stimulus appeared on the center screen. Once the experimenter observed that the infant looked at the center screen, the visual stimulus appeared on either the left or the right screen. When the infant turned their head to look at the side screen, the experimenter pressed a button and the auditory stimulus started playing. The trial ended when the infant looked away from the side screen for two consecutive seconds, or once the entire trial was complete (~ 21 s). If the infant looked to the side for less than 2 s total on any trial, that trial was skipped and then automatically repeated after the final trial.

Infants first completed two practice trials, which presented a non-language sound (Montréal: whistle sound; New Jersey: beep tones). Next, infants encountered 16 test trials. Half of the test trials were single-language trials, where all target words were in a single language. The other half were switched-language trials, where the language alternated between English and French (Montréal) or English and Spanish (New Jersey). Single-language trials were presented in a consistent language for each infant, to limit the occurrence of language switches between trials, such that language switches primarily occurred within switched-language trials. Monolingual infants heard the single-language trials in their native language, and bilingual infants were randomly assigned to hear single-language trials in their dominant or their non-dominant language.

The order of trials was pseudo-randomized, with the constraint that the same word pairs could not appear on consecutive trials, and no more than three trials of one type (single-language, switched-language) appeared consecutively. Each trial type appeared equally on the left and the right side, and no more than three consecutive trials appeared on the same side. The side of presentation was independent of the stimulus type.

Parents were asked about their child's language background using the LEQ structured interview and MAPLE approach (Byers-Heinlein et al., 2020), which asks about the child's lifetime exposure to different languages. Furthermore, parents filled out the Language Mixing Scale (Byers-Heinlein, 2013), which asks how often parents switch between their languages when speaking to their infant (see Appendix C). For one infant tested in Montréal and one tested in New Jersey, scores from the Language Mixing Scale were not available. As part of standard laboratory protocols, Montréal parents also completed a questionnaire with general demographic information (see Appendix D), as well as the MacArthur-Bates Communicative Development Inventories in American English (Fenson et al., 2007) and Quebec French (Trudeau et al., 1999). Questionnaires were completed either prior to or following the experimental portion of the study. Infants received a certificate and/or a small gift for their participation.

2.2.1.4 Coding

During the study, a trained experimenter blind to the auditory stimuli watched the infant through a live video feed and pressed buttons corresponding to a look to the left, right, or center monitor. A trained research assistant later re-coded the videos offline, frame-by-frame, with the sound off. For five participants, no offline coding was conducted, due to technical difficulties with the video recording ($n = 4$) and due to a procedural omission because the participant did not complete the minimum looking criterion ($n = 1$), and so the online coded data was analyzed. In these cases, we excluded trials that the experimenter flagged as coding errors, but retained the rest of the trials, as correlations between online and offline coding were high. We performed two checks to compare the experimenter's online coding to the offline coding (following B. Ferguson & Lew-Williams, 2016). First, the correlation between the offline and online coders' assessment of total looking times for each trial indicated high agreement for infants tested in Montréal (r = .98, 95% CI [.98, .99], t[430] = 110.03, $p < .001$) as well as those tested in New Jersey (r = .97, 95% CI [.95, .98], t[62] = 31.77, $p < .001$). Of particular importance for the head-turn preference procedure is whether the trial ended correctly during online coding, as the dependent measure is the total looking time for a trial. Recall that the experimental program ended a trial 2 s after an infant had looked away. A buffer of +/-.5 s was allowed to account for the time the online coder needed to react during the study. We thus examined the proportion of trials where the end time was either less than 1.5 s after the infant had looked away (indicating a trial that ended earlier than intended) or more than 2.5 s after the infant had looked away (indicating a trial that ended later than intended), according to offline coding. Overall, 10.89% of trials ended earlier than intended and 3.42% of trials ended later than intended. We did not exclude trials with these errors, but used offline-coded looking times in statistical analyses for all infants (when available).

2.2.2 Results

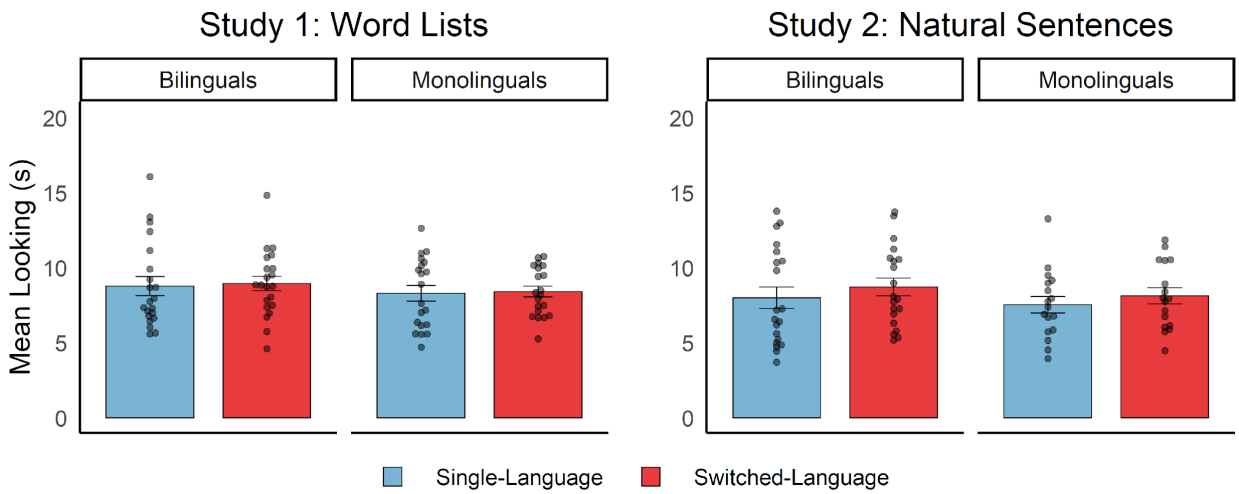
We assessed infants' detection of language switching by measuring their looking times (i.e., listening times) to single- vs. switched-language word lists. Looking times were log-transformed for all statistical analyses (Csibra et al., 2016), and figures show looking times prior to log transformation. After excluding trials with low looking time, the final sample of participants contributed on average 15.49 trials (range: 8–16). Data were analyzed using R version 4.0.3 (R Core Team, 2020) and the reproducible manuscript was created using papaja and citr (Aust, 2019; Aust & Barth, 2020). Looking times are shown in Table 2.3 and Figure 2.1 for each group (bilingual, monolingual) on each trial type (single-language, switched-language). Figure 2.2 shows the same data but using difference scores of looking time across trial types (calculated as $M_{switched}$ - M_{single}) in order to highlight individual infants' performance. The average looking time for single-language trials was 8.55 s (*SD* = 2.61 s) and for switched-language trials was 8.69 s (*SD* = 1.95 s).

Table 2.3 Mean Looking Times (Standard Deviations) and Effect Sizes in Studies 1 and 2

	Mean Looking Time (SD)		
Language Background	Single-Language	Switched-Language	Cohen's *d*
	Study 1		
bilingual	8.78 (2.88)	8.95 (2.23)	0.07
monolingual	8.30 (2.34)	8.42 (1.62)	0.06
	Study 2		
bilingual	8.01 (3.25)	8.73 (2.66)	0.24
monolingual	7.54 (2.29)	8.13 (2.20)	0.26

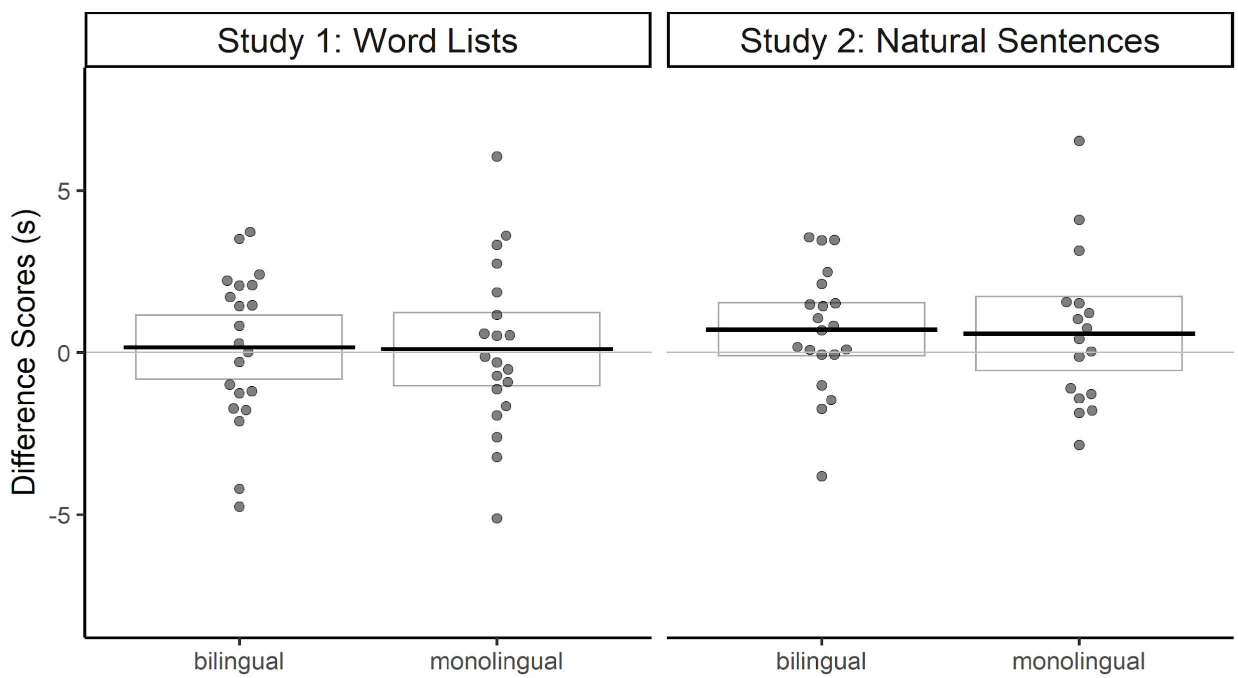
We conducted ANOVAs to investigate whether infants' looking patterns were affected by trial type (single-language, switched-language) and infants' language background (bilingual, monolingual). Due to difficulties in recruitment in New Jersey, we were unable to add test location as a between-subjects variable as we only had 7 participants in New Jersey. In the Supplementary Materials, we reported a separate ANOVA with the same design on data from infants tested in Montréal only, as well as descriptive data on the infants tested in New Jersey, which showed patterns consistent with the combined analysis presented here. In a 2 × 2 mixed ANOVA with trial type (single-language, switched-language) as a within-subjects factor and language background (bilingual, monolingual) as a between-subjects factor, we found no statistically significant main effects (trial type: $F[1, 39] = 0.21$, $MSE = 0.04$, $p = .652$, language background: $F[1, 39] = 1.13$, $MSE = 0.09$, $p = .295$) or interaction ($F[1, 39] = 0.01$, $MSE = 0.04$, $p = .930$). Thus, we did not find any evidence of greater attention to one trial type over the other, and bilinguals and monolinguals performed similarly.

Figure 2.1 Mean Looking Times for Single-Language and Switched-Language Trials



Note. Looking times averaged across participants are displayed separately for each language group. Left: Study 1 (word lists). Right: Study 2 (natural sentences). Error bars indicate standard error of the mean.

Figure 2.2 Difference Scores Showing Individual Participants' Looking Time to Switched-Language Relative to Single-Language Trials



Note. Higher values on the y-axis indicate longer looking to switched-language trials, and lower values indicate longer looking to single-language trials. This figure shows the same data as Figure 2.1, but displayed to highlight individual participant data.

Given the lack of a statistically significant difference, we then aimed to understand if our results reflected a true null effect or if we had insufficient power to detect a significant effect. To do so, we conducted equivalence tests (Lakens, Scheel, & Isager, 2018) to examine whether the effect size is likely smaller than our smallest effect size of interest, and therefore test for the absence of an effect. We set the smallest effect size of interest to be $d = 0.45$, based on the minimum detectable effect size we found in our sensitivity analysis (Lakens, 2017). The equivalence test comparing the observed effect size in Study 1 to the smallest effect size of interest (i.e., of $d = 0.45$) was significant ($t[40] = 2.42$, $p = 0.010$), meaning that the observed effect size ($d = 0.06$) was significantly within the interval from $d = -0.45$ to $d = 0.45$. This result indicates that we can reject a true effect size larger than $d = 0.45$ (or smaller than $d = - 0.45$). In other words, the equivalence test suggests a true null result for infants' looking times on single- and switched-language trials in lists of words.

2.2.2.1 Exploratory Analyses

We conducted additional analyses to examine individual differences in infants' detection of language switching. These analyses should be interpreted with caution, as they were exploratory. First, we calculated a Pearson correlation to examine whether infants who heard more language mixing at home performed differently in the experimental task than those who heard less language mixing at home. However, there was no evidence of a correlation between infants' language mixing score and their difference scores between looking time to single-language and switched-language trials ($r = .15$, 95% CI [-.17, .44], $t[37] = 0.93$, $p = .360$).

Second, we examined whether bilingual infants who heard single-language trials in their dominant language performed differently from those who heard the single-language trials in their non-dominant language. This was motivated by previous findings that infants' processing of language switching in word comprehension is asymmetrical across the two languages (Byers-Heinlein et al., 2017; Potter et al., 2018, 2019). Descriptively, the effect size for the difference between single and switched trials was larger for the subgroup of bilingual infants for whom the single-language trials were in their dominant language (dominant language: Cohen's $d = 0.38$, $n = 8$; non-dominant language: Cohen's $d = -0.09$, $n = 13$). However, the two-sample t-test comparing difference scores of looking to single and switched trials between participants tested in their dominant compared to non-dominant language was not statistically significant ($t[19] = 1.33$, $p = 0.198$). We note that the subgroups tested in each language were very small, and thus our analysis is both underpowered and subject to non-meaningful fluctuations in observed effect size. However, our results leave open the possibility that infants would detect language switches if the majority of words were in their dominant language, and this could be examined systematically in future investigations.

Third, we also explored whether bilingual infants who are exposed to their two languages in a more balanced manner (close to 50/50 %) respond differently in our task compared to unbalanced bilingual infants (those closer to 25/75 %). To investigate this, we correlated infants' performance in the task with their exposure to the non-dominant language (ranging from 25% or "unbalanced" to 50% or "balanced"). We found no statistically significant correlation (r = -.17, 95% CI [-.56, .28], t[19] = -0.75, p = .465), suggesting that balanced and unbalanced bilinguals performed similarly in Study 1.

2.2.3 Discussion

Study 1 indicated that 8- to 12-month-old bilingual and monolingual infants were unable to detect single-word language switches in word lists. There was no difference between infants' looking times on single-language compared to switched-language trials, and the effect size of the difference was statistically equivalent to zero. It may be that infants were unable to detect the language-switched words because the word lists used in the task are not typical of natural speech. Language input to children largely consists of multi-word utterances and sentences, rather than words in isolation (Brent & Siskind, 2001). Thus, naturally-occurring language switches may be more likely to be embedded within a sentence. Study 2 examined whether infants can detect language switches within natural sentences.

2.3 Study 2: Language Switches in Sentences

2.3.1 Method

2.3.1.1 Participants

Participants in Study 2 were 20 English–French or English–Spanish bilingual infants and 17 monolingual infants, largely overlapping with the infants who participated in Study 1 (see Supplementary Materials). Infants were tested in Study 1 and then Study 2 during the same visit. Inclusion criteria and number of infants excluded for health and language were the same as those in Study 1. One infant did not participate in Study 2 due to fussiness (their caregiver stopped participation during Study 1). Additionally, infants were excluded for equipment failure or experimenter error (n = 2), as well as for contributing fewer than 8 trials with at least 2 s of looking time (n = 9). The minimum looking time in Study 2 was set to 2 s, which was the first moment infants would be able to hear a language switch.

2.3.1.2 Stimuli

The visual stimulus was a video of a colourful animated pinwheel (Montréal) or an orange flashing light (New Jersey), distinct from but comparable in attractiveness and salience to the visual stimuli used in Study 1. Study 2 used the same nouns as Study 1, but they were presented in a sentence context instead of in isolation (see examples in Table 2.2 and the full list in Table E1 in Supplementary Materials). Both single- and switched-language trials began with a phrase in the same language, i.e., in English, French, or Spanish (e.g., English: "Do you like the …"/ "I want the …"), see Table E1 in the Supplementary Materials for the full list of stimuli. On single-language trials, the sentence-final target noun was presented in the same language as the sentence frame, whereas on switched-language trials the sentence-final noun was presented in the other language. The same bilingual female speakers produced the stimuli in Studies 1 and 2. The recordings used to create the stimuli for Study 2 were not spliced, but were recorded in a single session to retain the naturalistic articulatory features of the language switch. Care was taken to ensure that prosody, intonation, and phonetic realization were appropriate for the intended language, especially for words adjacent to the language switch. Words embedded in a statement sentence had a hill-shaped prosodic contour, and words embedded in a question had a rising prosodic contour (Nencheva et al., 2021). Trials lasted 20.4–24.6 s and were capped at the length of the shortest sound file, comparable to Study 1.

2.3.1.3 Procedure

The trial structure and experimental procedure were the same in Studies 1 and 2. Infants first completed Study 1, and then the researcher asked the parent to play with their infant while Study 2 was set up. This break was typically 2–3 minutes long, unless the infant needed to be fed or changed. We chose to present Study 2 after Study 1 because we expected that the naturalistic sentences presented in Study 2 would be more engaging to infants than the word lists presented in Study 1, thus increasing the likelihood of infants remaining attentive throughout both studies.

2.3.1.4 Coding

The same coding procedure from Study 1 was used for Study 2. For five participants, offline coding was not available, either due to technical problems with the video recording (n = 4), or due to a procedural omission because the participant did not complete the minimum looking criterion (*n* = 1). Data from the participants without video recording were retained because correlations between online and offline coding were high. The correlation between online and offline coding of total looking times for each trial was high for infants tested in Montréal ($r = .94$, 95% CI [.92, .95], $t[409] = 53.76$, $p < .001$) as well as New Jersey ($r = .95$, 95% CI [.87, .98], $t[14] = 11.69$, $p < .001$). Offline coding showed 9.51 % of trials ended earlier than intended and 6.28 % of trials ended later than intended. As in Study 1, we did not exclude trials with these errors, but used offline coded looking times when possible.

2.3.2 Results

The analytic strategy for Study 2 paralleled that of Study 1. After excluding trials with low looking and participants who subsequently did not provide the minimum number of trials to be included, the final sample on average contributed 15.08 trials (range: 9–16). The average looking time for single-language trials was 7.80 s ($SD = 2.82$ s) and for switched-language trials was 8.45 s ($SD = 2.44$ s). The results for Study 2 can be seen in the right panels of Figure 2.1 and Figure 2.2. Log-transformed looking times were entered into a 2 × 2 ANOVA with trial type (single-language, switched-language) as a within-subjects factor and language background (bilingual, monolingual) as a between-subjects factor. Trial type was not statistically significant (trial type: $F[1, 35] = 1.98$, $MSE = 0.04$, $p = .168$), and neither were the other main effects and interactions (language background: $F[1, 35] = 0.24$, $MSE = 0.18$, $p = .626$; interaction trial type × language background: $F[1, 35] = 0.04$, $MSE = 0.04$, $p = .840$). This indicates that neither bilingual nor monolingual infants detected differences between single-language and switched-language sentences.

As in Study 1, we used equivalence tests to test whether our observed effect size was smaller than our minimal detectable effect size of Cohen's $d = 0.45$. The equivalence test was not statistically significant ($t[36] = 1.31$, $p = 0.099$), meaning that the observed effect size ($d = 0.25$) in our study may not fall significantly within the interval of $d = -0.45$ to $d = 0.45$. This indicates that we cannot reject the hypothesis that the true effect size is Cohen's $d = 0.45$ or larger (or smaller than $d = -0.45$). Thus, our data are unsurprising in both cases, whether they were sampled from a null distribution centered around 0 (based on our ANOVA), or whether they were sampled from a distribution centered around $d = 0.45$ (based on the equivalence tests). In this case, we have insufficient data to draw a definite conclusion (Lakens et al., 2018), and future research will need to include a larger sample size.

2.3.2.1 Exploratory Analyses.

We again examined individual differences in performance as a function of exposure to language switching, and language dominance. There was no correlation between infants' language mixing score and their difference scores between looking time to single-language and switched-language trials (r = -.08, 95% CI [-.41, .26], t[33] = -0.48, p = .636). In the small subgroup of bilingual infants tested in their dominant language, anecdotally effect sizes were larger (dominant language: Cohen's d = 0.39, n = 9; non-dominant language: Cohen's d = 0.15, n = 11), although again, a two-sample t-test comparing difference scores for those tested in their dominant and non-dominant language was not statistically significant (t[18] = 0.25, p = 0.805), and subgroups tested in each language were very small. Thus, in Studies 1 and 2, descriptively larger effects were observed when single-language trials were presented in the dominant language, and this should be investigated more systematically. The correlation between balance of exposure for bilinguals and task performance trended towards but did not reach statistical significance (r = -.43, 95% CI [-.74, .01], t[18] = -2.04, p = .056), with less balanced children showing a numerically larger difference in looking to single- compared to switched-language trials. This finding should be replicated in a larger sample size before we can interpret it.

2.3.3 Comparison of Studies 1 and 2

We compared infants' looking times in Studies 1 and 2 in a 2 × 2 × 2 ANOVA with trial type (single-language, switched-language) and study (Study 1, Study 2) as within-subject factors, and language background as between-subject factor. We did not include the three-way interaction between trial type, study and language background due to concerns about low statistical power. Only participants who had usable data in both studies were included (n = 33). The main effect of study was statistically significant (F[1, 31] = 8.41, MSE = 0.06, p = .007), indicating that regardless of trial type, infants looked longer in Study 1 than Study 2 ($M_{Study\ 1}$ = 8.62, $M_{Study\ 2}$ = 8.12). None of the other main effects or interactions were statistically significant (trial type: F[1, 31] = 0.88, MSE = 0.06, p = .357; language background: F[1, 31] = 1.33, MSE = 0.19, p = .258; trial type × study: F[1, 32] = 0.88, MSE = 0.03, p = .356; study × language background: F[1, 31] = 0.01, MSE = 0.06, p = .905). This suggests that infants' looking times were not affected by trial type across both studies.

2.4 General Discussion

Bilingual infants regularly encounter switches between their two languages. Often these switches occur following long passages of speech, but they can also consist of a single word borrowed from the other language (e.g., "What a cute *chien* [fr. dog]"). Using the head-turn preference procedure, we tested whether bilingual and monolingual infants could detect single-word language switches by measuring their attention to single-language versus switched-language stimuli. We used two complementary types of statistical tests: ANOVAs in a null-hypothesis framework, and equivalence tests. ANOVAs indicate whether our observed effect is significantly different from zero and surprising under the null hypothesis, and equivalence tests indicate whether the observed effect is within the bounds of the smallest effect size of interest we set, and therefore practically not meaningful in the context of our study.

The results from Study 1 overall indicate that infants could not detect single-word language switches. However, the results from Study 2 were inconclusive: they do not provide strong evidence either way with regards to whether infants can detect switches that occur in naturalistic sentences. The effect size in Study 2 was small, but non-zero (d = 0.25). Such an effect size is not unexpected both in the case that infants cannot perform the task, and in the case that they can. Furthermore, we had expected language switches in sentences to be easier to detect, yet we did not find a difference between Studies 1 and 2 in the omnibus ANOVA.

There are two main reasons why we had expected a larger effect in Study 2. First, sentences are more typical of infants' everyday experiences, allowing them to engage mechanisms that support real-time language processing, such as statistical learning and prediction (see Potter & Lew-Williams, 2019). This may enable anticipation of the sounds and sound patterns that match the preceding words and 'surprisal' when the actual perceived sounds do not match those. Second, the switched-language sentences were produced naturalistically rather than by splicing, which might have afforded extra cues (e.g., coarticulation) in the speech signal even prior to the actual switch location (Fricke et al., 2016). In contrast, the isolated words were recorded in English-only and French-only contexts, and thus did not provide additional coarticulatory cues to the switch. Nevertheless, we believe that future studies should prioritize studying language switches in sentences, since the equivalence tests in Study 2 indicated a possibly meaningful effect size. At the same time, it should be noted that infants participated in our studies in a fixed order, to minimize fussiness and attrition as we predicted that Study 2 would be more interesting to the infants. A study published after the current work was undertaken suggests that infants with more previous experience with the head-turn preference procedure in other lab visits showed a larger novelty preference than those with less experience (see Santolin et al., 2021). Although in our case infants were tested twice within the same lab visit, we cannot rule out the possibility that infants' experience in Study 1 boosted their performance in Study 2.

We had predicted that bilinguals would show larger differences in response to language switching than monolinguals, but this was not supported by the data. Instead, neither bilinguals nor monolinguals showed evidence of detecting switches in either study. Within the bilingual group in Study 2 (language switches embedded in sentences), there was a non-significant trend for less balanced bilinguals to show a stronger effect than more balanced bilinguals, a finding which should be tested in a larger sample before it can be interpreted. Based on our results, we cannot say whether routine exposure to two languages affects infants' ability to detect deviations from a single language. Even if our results had shown that both bilinguals and monolinguals can detect language switches at this age, they may do so via different underlying processes. It may be that bilingual infants detect switches by recognizing certain sounds and words in each of their languages, which appear in different combinations in single- and switched-language trials. In contrast, monolingual infants were only familiar with one of the languages in the study, and thus could succeed on such a task by simply listening longer during trials that contained unfamiliar sounds and words (i.e., switched-language trials). Given our inconclusive results, our data cannot currently speak to these intriguing possibilities. To better understand the reasons behind bilingual and monolingual infants' performance on this task, future studies could avoid these issues of interpretation by using novel words that conform to phonological patterns in each language, but are unfamiliar to both groups.

Although we did not observe an impact of bilingual vs. monolingual language exposure in this investigation, there may be differential sensitivity to language switching depending on the frequency of switching in the bilingual input. This frequency is known to vary across different language communities and individual families (Bail et al., 2015; Byers-Heinlein, 2013; Kremin et al., in press). In a recent study on English–French bilingual infants in Montréal, parental language mixing was a relatively rare occurrence (Kremin et al., in press; Orena et al., 2019), but there may be much more language mixing in other communities, such as certain English–Spanish bilingual communities in the United States (Bail et al., 2015). There appears to be variation in how English–French and English–Spanish bilingual toddlers learn words in switched-language sentences (Byers-Heinlein et al., 2022), which raises the possibility that there are early-emerging differences in processing across different bilingual communities. While we had originally planned to compare these two populations in our analyses, the small sample size of the English–Spanish group prevented us from doing so. This is a limitation of our study that we hope will be remedied in future work. Studies that examine language mixing and bilingual input in different communities will be crucial for understanding pathways to bilingual proficiency. It would also be interesting to explore how individual differences in language mixing, language balance and dominant language status interact with the detection of language switches. We computed separate correlation measures for these factors and found no significant relationships, but we found a trend for enhanced sensitivity when detecting switches embedded in the dominant language. We were not able to directly compare the performance of French-English and Spanish-English bilinguals due to sample size limitations. However, we believe that including data from two bilingual communities is likely to make our results more generalizable. Future investigations exploring these moderators together would illuminate which aspects of the bilingual experience are most important to language switching.

Finally, it is important to consider whether our null/inconclusive results could be due to limitations of our experimental design to tap into infants' underlying ability. We tested infants using the head-turn preference procedure, which has yielded strong effect sizes in meta-analyses (Bergmann et al., 2018), and has been revealing about many aspects of infants' speech perception in infants of this age (e.g., Carbajal et al., 2021; Gasparini et al., 2021). However, we do note that other tasks have shown that older infants and toddlers are sensitive to single-word language switching. For example, several studies using both behavioural and ERP methods have found evidence that infants and toddlers respond distinctly to single-word language switching when viewing visual referents of the words spoken on the task (Byers-Heinlein et al., 2017; Kuipers & Thierry, 2012; Morini & Newman, 2019; Potter et al., 2018, 2019). Each of these studies tested children who were somewhat older than our participants (1.5 to 3 years old), who are more likely to have robust lexical representations of the stimuli. For such children, seeing a referent such as a picture of a dog when hearing "I like the …" could help infants generate predictions about the upcoming word, which then might help them recognize when they hear *chien* instead of dog. We cannot rule out that 8- to 12-month-old infants might succeed in a task that provides them with the visual referent. Nonetheless, our findings raise the possibility that perceptual information without referential context may be insufficient to support detection of single-word language switches in young infants.

Our overarching research question was how bilingual infants make sense of the dual language input they are exposed to. From the literature, we know that 0- to 4-month-olds can show successful cross-sentence language discrimination (for review, see Gasparini et al., 2021), and that in the second year of life they can detect single-word language switching (Byers-Heinlein et al., 2017; Kuipers & Thierry, 2012; Morini & Newman, 2019; Potter et al., 2018, 2019). Our study was novel because it tested whether infants can detect language switching during the period in-between these ages, when infants first start to recognize words. We did not find evidence that infants can detect single-word language switching at this age. This could mean that the ability to detect single-word language switching only starts to develop in the second year of life. If this is the case, it is unlikely that bilingual word learning in the context of mixed-language input relies on infants implicitly treating words from their two languages as belonging to different categories (Byers-Heinlein, 2014). It could be that more robust lexical knowledge is required for infants to be able to distinguish single words from two different languages. Since we did not find statistically significant results, we cannot resolve this question. We hope that the studies described here can be a stepping stone towards answering this question. Future research can also disentangle which cues are most relevant to children to detect single-word language switches. For example, we have discussed lexical knowledge and coarticulation, but other factors such as phonology and phonotactics can also play a role. Together, existing research reveals that language discrimination is multifaceted, spans multiple levels of language, combines multiple perceptual domains, and may change over the course of development. Future work will need to determine both the nature of 'successful' discrimination and the ultimate learning-related value of discriminating languages in the first place.

The present studies provide two main contributions. First, this work highlights a gap in the research into how bilingual infants make sense of their dual language input. While there is ample evidence that bilingual infants can discriminate some languages from birth when they hear whole sentences (Gasparini et al., 2021), it is only later in development around 20 months that the literature reports conclusive evidence for the ability to detect language switches for individual words (e.g., Byers-Heinlein et al., 2017). On the one hand, it may be that this is a much more difficult task with a protracted developmental time course. On the other hand, more sensitive experimental designs might be able to detect this ability earlier in development. This points to a more nuanced bilingual language development than previously thought. Second, we have reported results from two empirical studies aiming to directly test infants' detection of single-word language switches. The results from Study 1 were null, and those from Study 2 were inconclusive. Although these findings are somewhat unsatisfying, it is nonetheless important that they have a place in published literature rather than in the "file drawer" (Nelson et al., 2018). We have shared our materials, data, and analysis scripts, so that our research can contribute to future empirical studies, comprehensive reviews, and meta-analyses. We hope that this work will motivate further research into how infants navigate the dynamics of language input in environments where two languages are present, in turn illuminating how bilingual infants acquire proficiency in both their languages.

3 Keeping Track of Language: Can Bilingual and Monolingual Infants Associate a Speaker with the Language they Speak?

3.1 Introduction

Bilingual children regularly interact with individuals speaking different languages, but relatively little is understood about how the social aspects of daily interactions might support their language learning. For monolingual learners, learning about people and about language go hand in hand. For example, infants use information about what language a person speaks to reason about people's preferences and social affiliations (Kinzler, 2021). Similarly, bilingual adults use a person's language preference to generate predictions for language processing (Martin et al., 2016). For bilingual infants, being clued into who speaks what language could be especially useful. One theory of how bilingual infants acquire two languages and ultimately are able to speak each language is that they track the sounds, words and rules of their languages separately (Nazzi et al., 1998). In this case, being able to associate a person with the language they habitually speak could be helpful to encode information from the same language together (see also Kandhadai et al., 2014). In two studies each testing two age groups, we tested whether 5- to 18-month-olds spontaneously associate a person with the language they speak.

3.1.1 Bilingual Language Environments

Bilingual infants vary in terms of how they encounter their languages. One popular way to raise a bilingual infant is for two caregivers to each speak a different language to the child, also known as the "one-person-one-language" approach (Barron-Hauwaert, 2004). Early bilingual theorists hypothesized that this approach would allow children to separate their languages from the beginning, thus preventing confusion (Grammont, 1902). While more modern research has demonstrated that children can successfully acquire two languages from various family language configurations (De Houwer, 2007), it is nonetheless relevant to understand whether person-language associations are made in early bilingual language acquisition. While we are just beginning to describe bilingual infants' learning environments (Kremin et al., in press; Orena et al., 2020), it is probable that the people infants encounter use their languages in person-specific ways, and thus keeping track of who speaks what languages could make it easier for infants to detect structure in their language environment.

3.1.2 Foundational Skills

In order to form person-language associations, infants need to be able to discriminate between people and between the languages around them. Newborns can discriminate male and female voices (Floccia et al., 2000), and at least by 1 month of age, infants can discriminate unfamiliar faces when shown static images (de Haan et al., 2001). Same-gender voice discrimination seems to develop later, but starting at 7 months, monolingual infants can discriminate two female voices if the voices speak the infants' native language (Fecher & Johnson, 2018; Johnson et al., 2011). Bilingual infants may have an advantage in voice discrimination. In a set of studies, bilingual and monolingual 9-month-olds were habituated to a voice-face pairing, and at test the voice was changed. Bilingual infants were able to detect the change when they heard a familiar and an unfamiliar language, while monolingual infants only noticed the mismatch in a familiar language (Fecher & Johnson, 2019, 2022). Despite some of these nuances, overall, these studies show that infants learn to discriminate both same-gender as well as different-gender speakers early on, which is an important prerequisite for making person-language associations.

Infants can discriminate many language pairs from birth. This is well-documented in studies where infants hear long passages of speech and their reactions to a language switch are measured. Newborns' ability to discriminate languages depends on the similarity of the pairs of languages measured (Gasparini et al., 2021). Using auditory information only, infants can discriminate rhythmically different languages like English and Tagalog at birth (Byers-Heinlein et al., 2010), and rhythmically similar languages like Spanish and Catalan at least by 4 months of age (Bosch & Sebastian Galles, 2001; Molnar et al., 2014). Infants can also rely only on visual information to discriminate languages. In these studies, 4- to 8-month-old infants see a video of a person talking without sound, and are surprised when the person starts speaking a different language (Sebastian Galles et al., 2012; Weikum et al., 2007). When given both auditory and visual information, bilingual and monolingual 4-month-old infants can discriminate rhythmically distant languages, however bilingual infants do not seem to be able to discriminate rhythmically close languages (Birulés et al., 2018). Again, while there are some nuances to language discrimination, the ability to discriminate many language pairs emerges early.

3.1.3 Evidence for Sophisticated Reasoning about Language in Social Groups

While no studies have tested if infants can make spontaneous associations between a person and a language directly, there is evidence that infants are able to reason about how people they interact with use their language(s). This evidence comes from studies on mostly monolingual infants' reasoning about social groups. These studies differ from the studies reviewed so far in that they often test infants in highly social situations and/or measure preference for groups of people. However, these studies do highlight that language use can be a salient cue for reasoning about people. For example, monolingual 10-month-old infants preferred to take a toy from a person who spoke their native language (Kinzler et al., 2007). Monolingual infants also selectively imitate speakers of their own native language (Buttelmann et al., 2013; see also Howard et al., 2014). Moreover, monolingual infants use the languages they observe others speak to reason about their social relationships (Liberman et al., 2017). Based on these studies, we can conclude that even in the first year of life, monolingual infants make inferences about others based on the language these others speak, although similar studies have yet to be conducted with bilingual infants.

There is also evidence that infants take visual cues about a person into account when reasoning about what language an unfamiliar talker might use. At least by 11 months, monolingual English-exposed infants associated faces of East Asian women with hearing Cantonese (May et al., 2019), indicating that they use the specific language spoken as a cue to a person's identity in some cases. Further studies also show that monolingual infants take the speaker's race into account in language processing tasks (Singh et al., 2020; Weatherhead et al., 2021). These studies further underline that infants pay attention to the social context of language use.

3.1.4 Current Studies

We tested whether bilingual and monolingual infants spontaneously form an association between a person and the language they speak. Based on the studies of monolingual infants using language to reason about other people, as well as infants' perceptual abilities to discriminate languages and speakers, we expected that all infants would spontaneously form an association between a person and the language they speak. While there is evidence that bilingual adults form expectations about others' language use (Martin et al., 2016), this has not been tested in bilingual infants. The present paper will address this gap.

We tested infants using auditory-only stimuli in Study 1, and audiovisual stimuli in Study 2. The design of Study 1 was based on the previous literature on language discrimination, where most research uses auditory-only paradigms. In Study 2, we tested whether the addition of rich multimodal information changes infants' ability to learn person-language associations. We predicted that infants would have an easier time learning the person-language associations when exposed to audiovisual stimuli (Study 2), as the additional visual clues to the person's identity and language (Sebastian Galles et al., 2012; Weikum et al., 2007) might make it more engaging and salient to infants. However, we note that in a study reported after our data was collected, audiovisual stimuli possibly made it harder to detect a language switch for close language pairs (Birulés et al., 2018).

Each study used a familiarization-switch procedure. In the familiarization phase, infants encountered a woman speaking English and a man speaking French (or vice-versa). Then, infants were shown "Same" trials with a familiar speaker-language pairing (e.g., the woman speaking English) and "Switch" trials with a new speaker-language pairing (e.g., the woman speaking French). If infants are able to associate a speaker with their language, we expected them to look longer to the Switch than to the Same trial. We designed Study 2 such that we could also investigate pupil dilation as a measure of increased cognitive processing (Sirois & Brisson, 2014). Pupil dilation has been shown to be sensitive to language context (Byers-Heinlein et al., 2017; Fawcett, 2021). We predicted that infants' pupils would dilate on Switch trials compared to Same trials. This change in pupil size would indicate that infants learned the association between speaker and language in the familiarization, and therefore reacted differently to Switch and Same trials.

The age range tested in our studies spanned from 5 to 18 months. The age groups chosen represent three different developmental stages: before tuning into the sound structure of their native language (5 months), the early word learning phase (12 months), and the so-called vocabulary spurt phase (18 months). In Study 1, we collected full samples for two age groups: 5- and 12-month-olds. Based on the results of Study 1, in Study 2 we tested 12- and 18-month-olds, as we thought older infants might be more successful in our task. We expected that older infants would outperform younger infants, but we were also interested in seeing when infants would first succeed in our task.

We investigated the role of infants' language background by testing bilingual and monolingual infants. Bilinguals were learners of French and English, so they were familiar with both languages used in this study. Monolinguals had regular exposure to either English or French, and thus were familiar with only one of the two languages used. Our predictions for the role of language background were tentative. On the one hand, bilinguals might outperform monolinguals, as they are familiar with both languages and have regular opportunities to encode speaker-language pairings in their everyday life. On the other hand, bilingual infants might be used to a speaker switching to a new language, and in this case may not respond differently to Switch and Same trials. Language switches may be more obvious to detect for monolingual infants, for whom a speaker switching into or out of their native language is very salient as they hear few language switches in their daily life. Under this assumption, monolingual infants would perform better than bilingual infants.

Results from both studies showed that making person-language associations is difficult for infants, and that the explanation of how bilingual infants learn to discriminate and make sense of their social environment is potentially more complicated than expected.

3.2 Study 1: Auditory Cues Only

3.2.1 Method

This study was conducted in accordance with the Declaration of Helsinki, and was approved by the Human Research Ethics Board of Concordia University. Parents were informed of the study's purpose and procedure (see Appendix A for the consent form) and provided written consent prior to participation.

3.2.1.1 Participants

A total of 84 infants aged 5 or 12 months were included in the final sample. Each infant contributed data at a single age. Infants were recruited from government birth lists and a database of interested families in Montréal, Canada, a French-English bilingual city. Data collection took place between November 2013 and February 2017. When planning the study, we aimed for a sample of 16 per age and language group, which was a typical sample size at the time the study was planned (Oakes, 2017). We conducted a sensitivity analysis using G*Power (Faul et al., 2007) for a repeated measures ANOVA with a within-between interaction. The design has a power of 80% to detect an effect of at least $d = 0.51$. Due to a delay between data collection and data analysis, we tested more than 16 infants in all groups, and included all infants who were eligible based on our exclusion criteria. Infants were determined to be monolingual or bilingual based on their exposure to English and French.

Language exposure was measured via the Language Exposure Questionnaire (Bosch & Sebastian Galles, 2001) using the Multilingual Approach to Parent Language Estimates (MAPLE, see Measures for more detail, Byers-Heinlein et al., 2020). Bilingual participants were exposed to a minimum of 25% English and 25% French, and did not have more than 20% of exposure to a third language. The average exposure of bilingual infants to their dominant language was 62 % (range: 43-75%), and to their non-dominant language 36 % (range: 26-50 %). Nine bilingual infants also had exposure to a third language, at an average of 9 % (range: 2-19 %). Monolingual participants were exposed to either English or French at least 90% of the time for an average of 98% (range: 92-100 %). Bilingual infants' exposure to language mixing in their daily lives was measured by the Language Mixing Questionnaire (Byers-Heinlein, 2013). Language Mixing Questionnaire was not completed for 5 infants, those were treated as missing data. Bilingual infants scored on average 12.3 out of 30 (range: 0-30), with 30 indicating the highest level of language mixing.

To get to this final sample, 60 additional infants were tested but not included in the analysis due to failure to meet age criteria (*n* = 3), language criteria (*n* = 27), low birth weight (<2500 g) or premature birth (< 37 weeks of gestation, *n* = 3), reported major health issues (*n* = 1), and experimenter error (*n* = 1). Data for all remaining infants were analysed to check if they contributed enough data. For trials to be included into the final analysis, infants had to look for at least one second. For infants to be included in the analysis, they had to have completed at least one test trial of each type (Switch, Same). Following this procedure, 25 infants were excluded due to insufficient data, resulting in a final analyzed sample of 84 participants across the two age groups. Children's reported ethnicities were varied: the most reported category was European descent (*M* = 39 %), followed by parents listing multiple categories (*M* = 23%), including combinations of Afro-Canadian, Caribbean, and other identities, and a few parents used the write-in option to indicate they identified as either Canadian, Quebecois, or North American (*M* = 14%) or checked unknown (*M* = 12%). Furthermore, 6% indicated Afro-Canadian ethnicity, and all other options had less than 5% responses. Families had a high average level of maternal education (*M* = 15.9 years). A full breakdown of participants by age group and language group is displayed in Table 3.1.

We had also planned to test 9- and 18-month-old infants, and bilingual infants who hear *either* English or French as well as another language ("bilingual-other"), however due to recruitment difficulties these samples were not completed. Data from these infants were not analyzed for the purposes of this paper. For transparency, we have shared more details on these incomplete samples in the Supplemental Materials (Appendix F) and the full dataset is available on OSF (anonymized link for peer review: https://osf.io/psbxv/?view_only=a4fa383dfe7144c7bf0173e1f9c30a43).

Table 3.1 Descriptive Statistics by Age Group and Language Background

			Age (months)				
Age	Language group	N	Min.	Mean	Max.	% Female	% English dominant
			Study 1: Auditory-only				
5m	Bilinguals	19	5.0	5.5	6.0	63	53
	Monolinguals	19	4.8	5.7	6.5	42	42
12m	Bilinguals	24	11.7	12.3	13.0	54	67
	Monolinguals	22	11.8	12.5	3.3	36	41
			Study 2: Audiovisual				
12m	Bilinguals	18	11.8	12.5	13.2	33	67
	Monolinguals	24	11.8	12.6	13.2	42	67
18m	Bilinguals	21	17.7	18.5	19.2	67	52
	Monolinguals	18	17.8	18.5	19.1	50	56

Note. For monolinguals, the column % English dominant denotes the percent of children who are English monolinguals, the remainder of monolinguals are French-dominant.

3.2.1.2 Stimuli

Auditory stimuli consisted of voice recordings made by one male and one female English-French bilingual speaker. Both speakers had acquired Canadian English and Quebec French from birth and spoke with a Montréal-native accent in either language as judged by a different highly proficient English-French bilingual. The passages that were recorded for the current study were excerpts from *The Little Prince*, which was chosen as it was used in Weikum et al. (2007). The speakers read the passages in adult-directed speech, as most of the studies on language discrimination have used adult-directed speech (Gasparini et al., 2021). For each trial, we selected sentences read by a single speaker, such that pitch and duration were matched across speakers and languages. All recordings were normalized to 70dB using PRAAT (Boersma & Weenink, 2016) and each trial lasted approximately 18 seconds. A transcript of the passages used and stimuli files can be accessed on the OSF platform (https://osf.io/psbxv/?view_only=a4fa383dfe7144c7bf0173e1f9c30a43). The visual stimulus was constant throughout the study, and consisted of a picture of a field of flowers (see Fig. 1) which was chosen to provide a visually pleasing stimulus for infants to look at.

3.2.1.3 Measures

Demographic information. Parents completed a demographics questionnaire that asked about infants' health history, gestational age, birth weight, family structure, and parents' level of education (see Appendix D).

Multilingual Approach to Parent Language Estimates (MAPLE). An adapted version of the Language Exposure Questionnaire (Bosch & Sebastian Galles, 2001) was administered as a structured interview using MAPLE (Byers-Heinlein et al., 2020). The questionnaire can be found in Appendix B. The experimenter asked questions about language exposure on a typical day for each month in the child's life and then calculated the percentage of the child's exposure to English, French, and other language(s). This information was used to calculate the percentage of time that infants had been exposed to each language since birth.

MacArthur-Bates Communicative Development Inventory (MCDI). In order to assess infants' vocabulary size for comprehension and production, parents of children 12 months or older filled out either the MCDI – Words and Gestures (12-month group) or MCDI – Words and Sentences (18-month group). Parents of monolingual infants completed the MCDI in its original English version (Fenson et al., 2007) or in its Canadian French in Québec adaptation (Trudeau et al., 1999) as applicable, while parents of bilingual infants completed both versions.

Language Mixing Questionnaire. Parents completed a self-report questionnaire (see Appendix C) that assessed the frequency of language mixing when speaking directly to their child, including code switching and borrowing (Byers-Heinlein, 2013). This measure was only collected for bilingual infants.

3.2.1.4 Procedure

Participants were tested in a dimly lit sound-attenuated room. Infants sat on their parent's lap facing the screen of a Tobii T60-XL eye tracker (Tobii® Technology, 2010) that gathered eye-gaze data. Parents wore darkened sunglasses and listened to music on headphones to minimize their interaction with the infants during the study. The experimenter controlled the study from an adjacent room using Tobii Studio software and was blind to the trial type presented. Before stimuli presentation, the eye-tracker was calibrated to the participants' eyes using a five-point infant calibration routine. A bouncing circle changing in colour and shape was presented before each trial to redirect the infant's gaze to the centre of the screen.

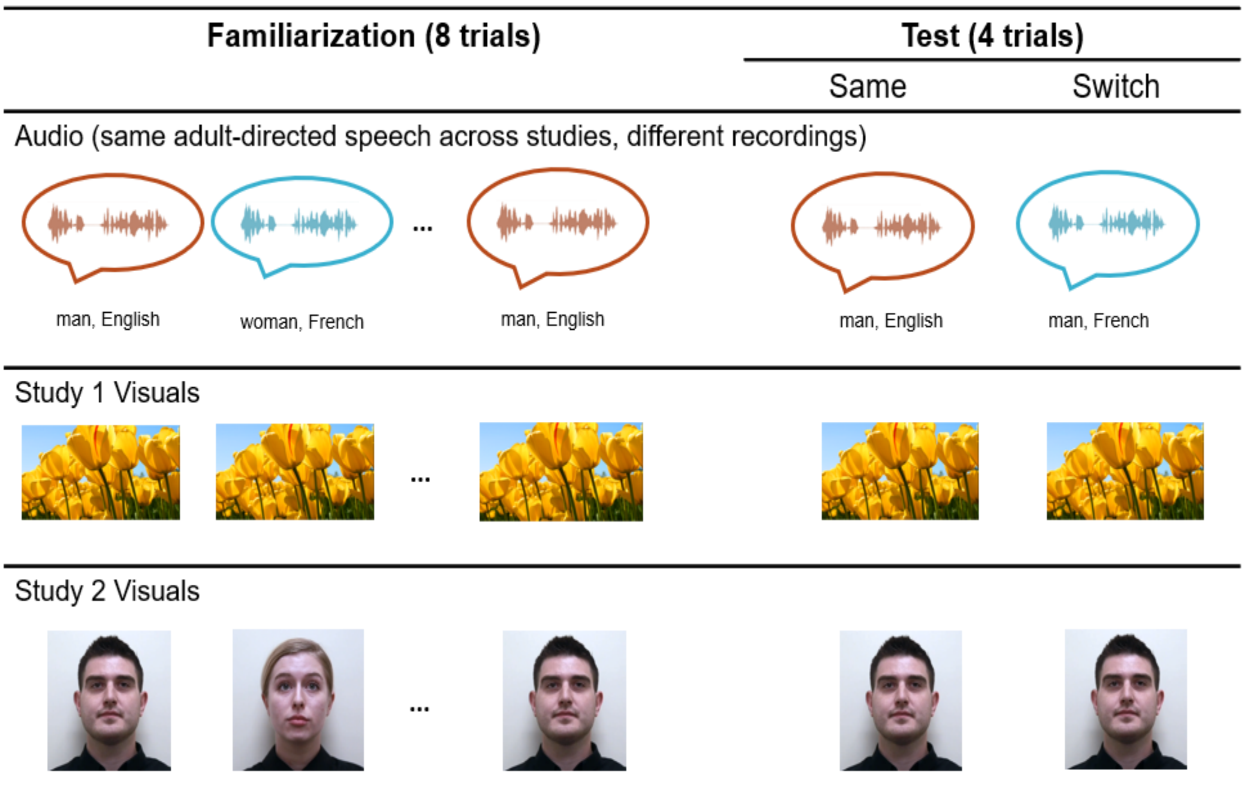
Infants were familiarized to a total of 8 trials of either the male or the female speaker talking, interleaved in a pseudorandomized order. Each trial used a different passage from the story *The Little Prince*. Infants heard a consistent speaker-language pairing, e.g., the female voice speaking English and the male voice speaking French. The speaker-language pairing was counterbalanced across infants, such that some infants heard the female speaker speaking French and the male speaker speaking English and vice versa. In the test phase, infants heard two trials with the same speaker-language pairing as in familiarization (Same trials) and two trials with each of the familiarization speakers talking in a different language (Switch trials). All infants saw four test trials, one of each trial type per speaker. The order of test trials was also counterbalanced, such that some infants heard a Switch trial first and some heard a Same trial, and some heard the male speaker first and some heard the female speaker first. An example of a study order is shown in Figure 3.1. The eight familiarization trials within each order were presented quasi-randomly in one of eight experimental orders, such that no more than two trials of the same speaker-language pairing appeared in a row. The experiment lasted approximately 5 minutes. After completion of the study, infants received a participation certificate and a small gift.

3.2.2 Results

3.2.2.1 Analytic Approach

The area of interest was defined as the entire display area (1920×1200). The dependent variable was total looking time in seconds on each test trial, which we log-transformed following recommendations for infant looking time studies (Csibra et al., 2016). As a robustness check we repeated all looking time analyses with the untransformed looking times, and report the full results tables in the Supplemental Materials as well as a short summary in the sections below. As described in the Participants section, infants who looked for less than 1 s on more than two of the four test trials were excluded from analysis, and following that, infants who did not contribute at least one type of each trial (Switch, Same) were excluded. The remaining 5-month-olds completed on average 12.0 trials with 6.41s of looking time per trial, and 12-month-olds completed on average 11.9 trials with an average looking time of 4.01s per trial.

Figure 3.1 Example of one Trial Order for Studies 1 and 2

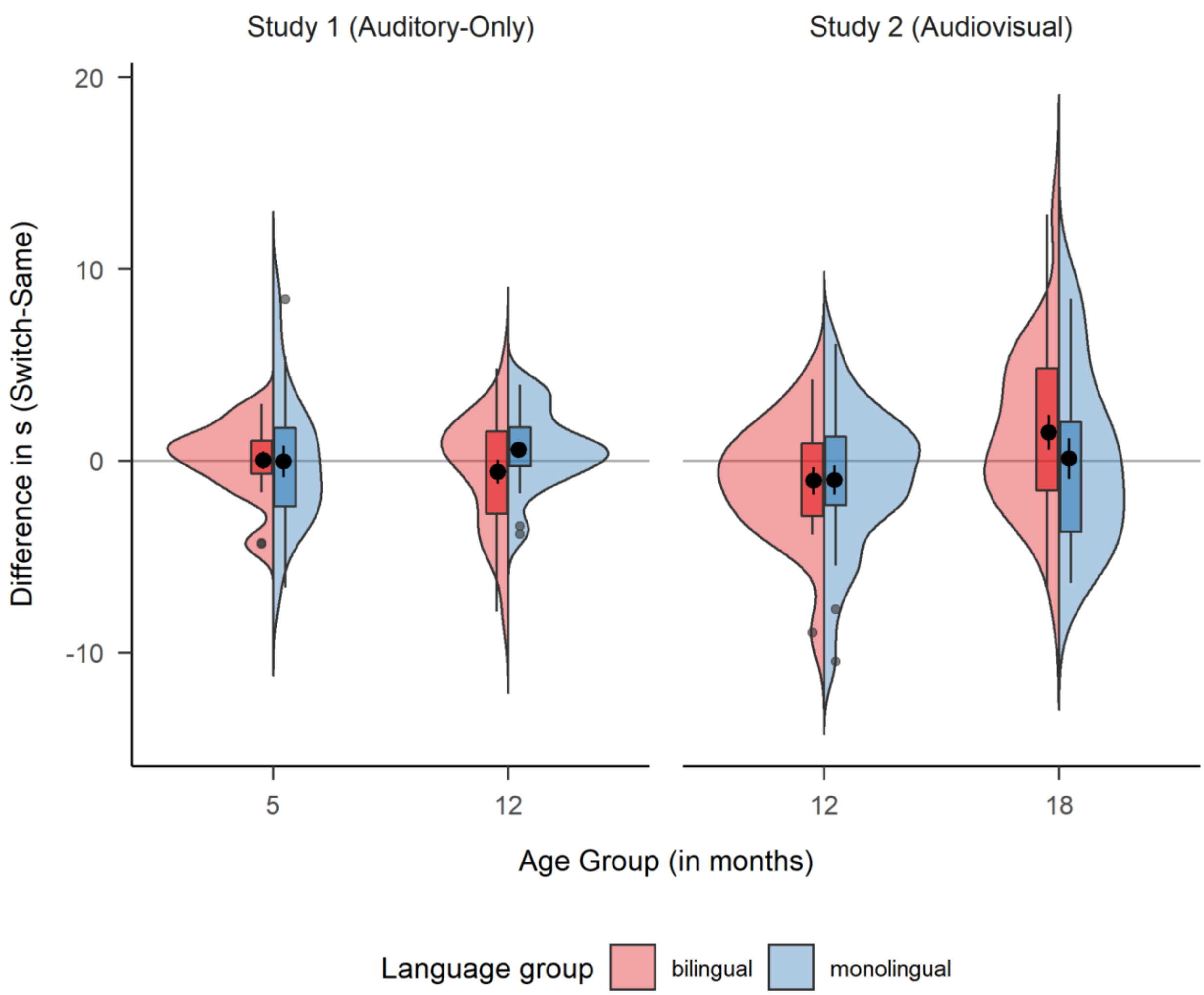


Note: This figure shows two of four test trials to illustrate, in the remaining test trials the woman speaks either the same or a different language. Trial order was pseudorandomized, and assignment of speaker and language combination was counterbalanced across infants. The outline colour of the speech bubble represents the language spoken, English (red) or French (blue).

3.2.2.2 ANOVA and Equivalence Tests

A 2 (language background: monolingual, bilingual) × 2 (test trial type: Same, Switch) mixed factorial ANOVA was conducted separately for each age group. Language background was a between-subjects factor and trial type was a within-subjects factor. The main dependent variable was log-transformed looking times, and note the same pattern of results was found in the robustness analysis (Table F1 in Appendix F). Difference scores of looking time in Switch and Same trials can be seen in Figure 3.2 (left panel).

Figure 3.2 Difference Scores for Looking to Switch versus Same by Study and Age Group



Note. Black dots with whiskers inside the boxplots show the mean and the standard error of the mean. Violin plots show the distribution of values.

For the 5-month-olds, we found no effect of trial type ($F[1, 36] < 0.01$, $MSE = 0.13$, $p = .955$, $\eta^2_p < .01$), indicating that infants did not look differently to Switch trials ($M = 6.42$) than to Same ($M = 6.41$, Cohen's $d < 0.01$). There was no effect of language background or interaction between language background and trial type (language background: *F*[1, 36] = 0.26, *MSE* = 0.57, *p* = .615, η^2_p = .007; interaction: *F*[1, 36] < 0.01, *MSE* = 0.13, *p* = .981, η^2_p < .01) in the 5-month old group.

Similarly, for the 12-month-olds, we also found no effect of trial type (*F*[1, 44] = 0.04, *MSE* = 0.26, *p* = .842, η^2_p = .001), indicating no difference in looking time to Switch (*M* = 3.88) and Same trials (*M* = 4.13, Cohen's *d* = - 0.10). There was also no effect of language background or interaction with language background (language background: *F*[1, 44] = 0.36, *MSE* = 0.57, *p* = .554, η^2_p = .008; interaction: *F*[1, 44] = 1.04, *MSE* = 0.26, *p* = .313, η^2_p = .023).

To summarize, across 5- and 12-month-old infants, there was no evidence that infants noticed when a speaker switched languages. Infants' prior language exposure did not seem to affect their behaviour.

We used equivalence tests to test if there may be a true effect of trial type that our analysis was underpowered to detect. Unlike null-hypothesis tests, which can only conclude whether two conditions are significantly different, equivalence tests allow us to conclude that two conditions are statistically equivalent. Equivalence tests necessitate choosing the smallest effect size of interest, which we set to be the smallest effect size we can detect with our planned sample size based on our sensitivity test, *d* = 0.51. Equivalence tests indicated that proportional looking time to Switch and Same trials was statistically equivalent, for both 5-month-olds (*t*[37] = 3.10, *p* = 0.002) and 12-month-olds (*t*[45] = 3.27, *p* = 0.001). The same pattern of results was found in the robustness analysis using untransformed looking times, and can be seen in Table F2. This indicates that we can reject the hypothesis that the true effect size is larger than *d* = 0.51 (or smaller than *d* = -0.51). We thus conclude that infants' equal looking times on auditory-only Same and Switch trials are likely a true null result.

3.2.2.3 Exploratory Analyses

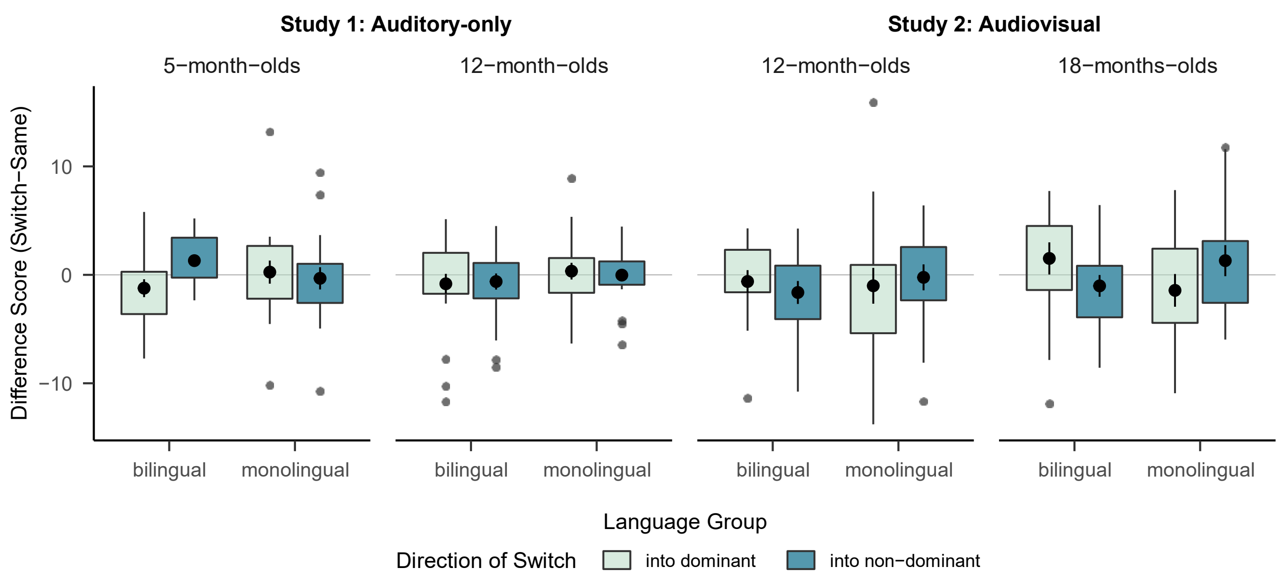
We also explored associations between the infants' language exposure and their task performance. Because we did not plan for these analyses, the results should be treated with caution. First, we correlated infants' exposure to their dominant language (native language for monolinguals) and their difference score (Switch - Same trials) to see if more exposure to one language affects the size of the difference. There was no significant correlation for 5-month-olds (r = -.04, *95% CI* [-.35, .29], *t*[36] = -0.22, *p* = .831) or 12-month-olds (r = .18, *95% CI* [-.12, .45], *t*[44] = 1.22, *p* = .229), suggesting that the amount of exposure was not associated with differential performance in our auditory-only task. The same pattern of results was found in the robustness analysis (Table F3 in Appendix F).

We computed a correlation between the parental report on the amount of language mixing they use with their child and differential looking time to Switch and Same trials, to test whether more exposure to language mixing would help infants make the association between a language and a person. Language Mixing scores were only collected for bilinguals, and thus the correlation was computed for the 19 5-month-olds and 24 12-month-olds for which these scores were available. We found no correlation between language mixing and differential looking time to Switch and Same trials in 5-month-olds (r = -.09, *95% CI* [-.53, .40], t[16] = -0.35, p = .730) or 12-month-olds (r = .06, *95% CI* [-.39, .49], t[18] = 0.25, p = .804). The same pattern of results was found in the robustness analysis (Table F4 in Appendix F). These results indicate that more exposure to language mixing was not associated with better performance in the auditory-only task.

We also tested whether the direction of the language switch affected performance. For each infant, one Switch trial was a switch into their dominant language (native language for monolinguals), and the other was a switch into their non-dominant language. To investigate whether the direction of the switch affected infants' performance, we calculated an ANOVA with the difference in looking time (Switch-Same) as the dependent variable, and the direction of switch (into dominant, into non-dominant) as a within-subjects variable, and language group (bilingual, monolingual) as a between-subjects variable. Only infants who completed all four test trials were included: 38 5-month-olds (19 bilingual, 19 monolingual) and 44 12-month-olds (22 bilingual, 22 monolingual). Difference scores by language group and direction of switch can also be seen in the left two panels of Figure 3.3.

In the 5-month-old group, we found no effect of direction of switch (F[1, 36] = 1.34, MSE = 1.76, p = .255, η^2_p = .036), no effect of language background (F[1, 36] = 0.09, MSE = 1.24, p = .772, η^2_p = .002), and no interaction between the two variables (F[1, 36] = 1.61, MSE = 1.76, p = .212, η^2_p = .043). Similarly, in the ANOVA for the 12-month-old group, we found no effect of direction of switch (F[1, 42] = 0.02, MSE = 20.18, p = .897, η^2_p < .001), as well as no effect of language background (F[1, 42] < 0.01, MSE = 22.18, p = .951, η^2_p < .001) or interaction between the two variables (F[1, 42] = 0.49, MSE = 20.18, p = .488, η^2_p = .012). For both analyses, robustness analyses showed the same pattern (see Table F5 in Appendix F). Taken together, there was no evidence that 5- or 12-month-olds are affected by whether the speaker switches into or out of their dominant language when hearing a person switching languages.

Figure 3.3 Difference Scores of Looking to Switch versus Same, by Language Group and Direction of Switch



Note. Black dots with whiskers inside the boxplots show the mean and the standard error of the mean.

3.2.3 Discussion

In Study 1, bilingual and monolingual infants aged 5- and 12-months-old did not show evidence of forming person-language associations. One reason why infants may not have formed these language-person associations could be because they did not have access to the facial cues associated with seeing someone speak that they would have in real world interactions. These cues have been shown to allow infants to detect a language switch in a visual-only task (Sebastian Galles et al., 2012; Weikum et al., 2007). Furthermore, adding visual information might make the speaker's identity more salient and might therefore facilitate the infants' learning about the speaker. Thus, Study 2 examined whether infants could form person-language associations when both seeing and hearing a speaker. In addition to measuring looking times, we also designed this study such that we can measure pupil dilation, by having the onset of speech start after the onset of the visual stimulus, to be able to measure any effects of cognitive processing in pupil dilation separately from the luminance changes associated with the onset of the visual stimulus. Together, looking times and pupil dilation allow us to test whether infants learned the association between a person and the language they speak in an audiovisual task.

3.3 Study 2: Audiovisual Cues

3.3.1 Method

The procedures for Studies 1 and 2 were identical, except for the video stimuli used in Study 2. Differences in protocols are outlined below.

3.3.1.1 Participants

A total of 81 bilingual or monolingual infants were included in the final sample. We recruited a 12-month-old group and an 18-month-old group; 5-month-olds were not tested as this population was harder to recruit than slightly older infants and did not show significant effects in Study 1. All other criteria for participation were identical to Study 1. Data collection took place between September 2017 and September 2019. We planned to test 16 infants per group to match the samples for Study 1, and for all groups we ended up with a slightly larger group of participants due to delays between data analysis and recruitment, as in Study 1. Twenty-seven additional infants were tested but not included in the analysis due to failure to meet age criteria (n = 1), pre-established language criteria (n = 20), or low birth weight or premature birth (n = 6). As in Study 1, data for the remaining infants were analysed to check if they contributed enough data. Following this procedure, eight infants were excluded, resulting in a final analyzed sample of 81 infants across the two age groups. More information on the samples can be seen in Table 3.1.

As in Study 1, language exposure was measured using a structured interview (Bosch & Sebastian Galles, 2001; Byers-Heinlein et al., 2020). Bilingual infants' average exposure to their dominant language was 62 % (range: 44-75 %), and to their non-dominant language 35 % (range: 25-49 %). Nine bilinguals had some exposure to a third language, on average that exposure was 13 % (range: 2-19%). Monolingual infants were exposed to an average of 98% of either English or French (range: 90-100%). Parental responses on the Language Mixing Questionnaire (Byers-Heinlein, 2013) were collected for bilingual infants, as in Study 1. Bilingual infants scored on average 10.0 out of 30 (range: 1-24), with 30 indicating the highest level of language mixing (for one infant, the questionnaire was not completed, this data point was treated as missing).

The reported ethnicities of our sample were as follows: most reported European descent (51%), followed by multiple responses (33%), including Afro-Canadian and Caribbean. All other categories had less than 5% responses. Maternal education was on average 16.4 years.

3.3.1.2 Stimuli

Audiovisual stimuli consisted of video recordings of one male and one female English-French bilingual speaker reading the same passages used in Study 1. As the two original speakers recorded in Study 1 were not available, we recorded two different early learners of Canadian English and Quebec French. Each speaker was filmed in front of a white background while wearing a black shirt. Speakers were instructed to keep a neutral face and to move their head as little as possible. Each video started with ~2 s of silence before speech onset, where the speaker looked straight at the camera. This was necessary for analysing the pupil dilation at the speech onset, to avoid conflating the onset of the video and the corresponding luminance changes with the onset of speech effects. We used the same passages from *The Little Prince* as in Study 1. The passages were read in adult-directed speech and each one lasted approximately 20 s. As in Study 1, each video clip consisted of a single speaker reading a single passage. The audio of all recordings was normalized to 70dB using PRAAT and the videos' white balance was adjusted using iMovie. An example of the audiovisual stimuli is shown in Figure 3.1, and all stimuli files are also available on the OSF platform (anonymized link for peer review: https://osf.io/psbxv/?view_only=a4fa383dfe7144c7bf0173e1f9c30a43).

3.3.1.3 Measures and Procedure

Parents completed the same questionnaires as in Study 1. The same familiarization-test paradigm and design as in Study 1 was used.

3.3.2 Results

3.3.2.1 Analytic Approach

The analytic approach for looking time data was parallel to Study 1. Once again, our main analysis focused on log-transformed looking times, and we conducted robustness analyses on raw looking times. Additionally, we were also able to analyze pupil dilation data. This was possible in this study but not in Study 1 because we designed the stimuli such that the speech onset (which differentiates Switch and Same trials) started after the infants' pupils had adjusted to the change in luminance associated with the onset of the video stimulus. Twelve-month-olds on average completed 11.6 trials with a mean looking time of 7.96 s per trial, and 18-month-olds completed 11.5 trials and had a mean looking time of 9.32 s per trial.

3.3.2.2 ANOVA and Equivalence Tests

As in Study 1, for each age group, we calculated separate 2 × 2 ANOVAs with trial type (Same, Switch) as a within-subjects factor and language background (bilingual, monolingual) as a between-subjects factor. In the 12-month-old group, effect of trial type trended towards but did not reach statistical significance ($F[1, 40] = 3.89$, $MSE = 0.09$, $p = .056$, $\eta^2_p = .089$). In the robustness analysis on untransformed looking time data, this effect was more clearly not statistically significant ($p = .124$, $\eta^2_p = .058$, see also Table F1 in Appendix F). Twelve-month-olds looked numerically shorter looking to Switch trials ($M = 7.58$ than Same ($M = 8.34$), Cohen's $d = -0.22$). The effect of language group trended towards but did not reach statistical significance ($F[1, 40] = 3.53$, $MSE = 0.39$, $p = .067$, $\eta^2_p = .081$). Numerically, bilingual infants ($M = 6.99$s) looked less long than monolingual infants ($M = 8.69$s, Cohen's $d = -0.49$), regardless of trial type. There was no interaction between language group and trial type ($F[1, 40] = 0.48$, $MSE = 0.09$, $p = .491$, $\eta^2_p < .012$). The pattern of results in the robustness analysis was the same.

In the 18-month-old group, we found no effect of trial type ($F[1, 37] = 1.76$, $MSE = 0.24$, $p = .193$, $\eta^2_p = .045$), as evidenced by similar looking times on Switch ($M = 9.72$s) and Same trials ($M = 8.91$s, Cohen's $d = 0.16$). Again, there were also no effects of language group ($F[1, 37] = 1.18$, $MSE = 0.66$, $p = .284$, $\eta^2_p = .031$) or interaction between language group and trial type ($F[1, 37] = 0.49$, $MSE = 0.24$, $p = .488$, $\eta^2_p = .013$). The pattern of results in the robustness analysis on untransformed looking time data was the same. While we did find a marginal difference between Switch and Same trials in the 12-month-olds, the direction was opposite to the expected effect, and different to the direction of the effect in 18-month-olds. We thus speculate that the marginal effect found for 12-month-old is likely due to chance.

We used equivalence tests to determine whether infants' looking times for Switch and Same trials were statistically equivalent. As in Study 1, we used $d = 0.51$ as our smallest effect size of interest, which is the smallest effect size our planned sample has 80% power to detect. In the 12-month-old group, the equivalence tests were not significant ($t[41] = -1.33$, $p = 0.095$), which means we cannot reject the hypothesis that the true effect is larger than $d = 0.51$ (or smaller than $d = –0.51$). However, in the robustness analysis on untransformed looking time, the equivalence test was significant ($t[41] = -1.73$, $p = 0.046$, see also Table F2 in Appendix F).

These conflicting findings in conjunction with the marginal but not statistically significant main effect of trial type in the ANOVA, indicate that our observed data are unsurprising in either case, whether they were sampled from a distribution around 0 (as tested in the ANOVA) or a distribution around $d = 0.51$ (as tested in the equivalence test). For the 18-month-olds, the equivalence tests were significant ($t[38] = 1.86$, $p = 0.035$), meaning that for 18-month-olds we can reject the hypothesis that the effect size is larger than $d = 0.51$ (or smaller than $d = -0.51$). The results of the robustness analysis on untransformed data showed the same result (see Table F2 in Appendix F). In summary, there may be a smaller-than-expected effect in the opposite direction than predicted in 12-month-olds, which would need to be explored in future research with a larger sample size. In the 18-month-olds sample, we found equal looking times to Switch and Same trials based on a smallest meaningful effect size of $d = 0.51$. Taken together, there is no evidence that 18-month-olds are affected by whether the speaker switches into or out of their dominant language when observing a person switching languages, and inconclusive evidence for 12-month-olds, which necessitates future research with a larger sample size.

3.3.2.3 Exploratory Analyses

As in Study 1, we conducted some exploratory analyses that should be interpreted with caution, as they were not planned when designing the study. We found no correlation between the exposure to the dominant language (native language for monolinguals) and differential looking time to Switch and Same trials in 12-month-olds ($r = .08$, *95% CI* [-.23, .38], $t[40] = 0.53$, $p = .599$) or 18-month-olds ($r = -.15$, *95% CI* [-.44, .18], $t[37] = -0.90$, $p = .376$). The same pattern of results was found in the robustness analysis, see Table F3 in Appendix F. This indicates that more exposure to the dominant language did not help infants detect when a person switched languages.

We also computed a correlation between the parental reports on the amount of language mixing parents use with their child and differential looking time to Switch and Same trials, to test whether more exposure to language mixing would help infants make the association between a language and a person. Language Mixing scores were only collected for bilinguals, and thus the correlation was computed for the 18 12-month-olds and 21 18-month-olds for which these scores were available. We found no correlation between language mixing and differential looking time to Switch and Same trials in 12-month-olds ($r = -.09$, *95% CI* [-.53, .40], $t[16] = -0.35$, $p = .730$) or 18-month-olds ($r = .06$, *95% CI* [-.39, .49], $t[18] = 0.25$, $p = .804$). The same pattern of results was found in the robustness analysis, see Table F4 in Appendix F. Thus, more exposure to language mixing did not help infants make an association between a person and the language that person speaks.

To explore how the direction of the switch in the test trials affected infants' performance, we again calculated a 2 × 2 ANOVA using log-transformed differential looking time (Switch-Same trials) as the dependent variable. Direction of switch (into dominant language, into non-dominant language) was a within-subjects variable, and language background (bilingual, monolingual) was a between-subjects variable. As in Study 1, we only included infants who completed all four test trials, resulting in 32 12-month-olds ($n_{\text{bilingual}} = 15$, $n_{\text{monolingual}} = 17$) and 29 18-month-olds ($n_{\text{bilingual}} = 15$, $n_{\text{monolingual}} = 14$). In the 12-month-old group, we found no statistically significant effect of direction of switch ($F[1, 30] < 0.01$, $MSE = 0.68$, $p = .959$, $\eta^2_p < .001$) or language background ($F[1, 30] = 0.81$, $MSE = 0.36$, $p = .375$, $\eta^2_p = .026$), or their interaction ($F[1, 30] = 1.19$, $MSE = 0.68$, $p = .284$, $\eta^2_p = .038$). In the 18-month-old group, we found no main effect of direction of switch ($F[1, 27] = 0.12$, $MSE = 0.37$, $p = .729$, $\eta^2_p = .005$) and no main effect of language background ($F[1, 27] < 0.01$, $MSE = 0.63$, $p = .994$, $\eta^2_p < .001$). The interaction between direction of switch and language background was statistically significant ($F[1, 27] = 4.96$, $MSE = 0.37$, $p = .034$, $\eta^2_p = .155$). Robustness analysis on untransformed looking times showed the same pattern of results, and can be seen in Table F5 in the Supplemental Materials (Appendix F). The interaction can also be seen in the right two panels in Figure 3.3. Bilingual infants looked longer to Switch trials when the switch was into the dominant language compared to when it was into the nondominant language, while the opposite pattern was observed for monolingual infants. This pattern would necessitate replication to test if this is a true effect.

3.3.2.4 Pupil Dilation

We investigated pupil dilation as a measure of increased cognitive processing in response to Switch compared to Same trials. We followed recommendations for analyzing pupillometry data in our data preparation (Forbes, 2020; Sirois, 2018) . In cases where pupil data was missing for only one eye, we regressed the values from the available eye since pupil dilation measures for both eyes were highly correlated ($r = .94$). We then averaged pupil size across both eyes, and down sampled the data to 200ms time bins. We removed trials with less than 25% of valid pupil size measurements, which was 65 trials (21.8% of all trials). We then removed infants who did not contribute at least one trial of each type (Same, Switch), which removed 20 participants (24.7% of the sample). This resulted in 31 12-month-olds (12 bilinguals, 19 monolinguals) and 30 18-month-olds (16 bilinguals, 14 monolinguals). We then filtered data using a median filter to remove noise, and interpolated across missing data such as blinks using linear interpolation. Finally, we set the baseline to the 200ms time window before onset of speech, and subtracted the mean pupil size during the baseline from the within-trial pupil sizes to get change in pupil size from baseline.

We then computed separate 2 × 2 ANOVAs for each age group, with pupil dilation as dependent variable, and trial type (Switch, Same) as a within-subjects variable and language background (bilingual, monolingual) as between-subjects variable. In 12-month-olds, we found no effect of trial type (F[1, 29] = 0.01, $MSE < 0.01$, $p = .909$, $\eta^2_p < .001$), as evidenced also by similar change in pupil size for Switch ($M = 0.085$) and Same trials ($M = 0.084$) . We also did not find an effect of language background (F[1, 29] = 0.02, $MSE = 0.02$, $p = .894$, $\eta^2_p = .001$) or an interaction between language background and trial type (F[1, 29] = 1.03, $MSE < 0.01$, $p = .317$, $\eta^2_p = .034$). In 18-month-olds, we found no effect of trial type (F[1, 28] = 1.13, $MSE = 0.01$, $p = .297$, $\eta^2_p = .039$), as evidenced by similar pupil dilation on Switch ($M = 0.110$) and Same trials ($M = 0.088$) . We further found no main effect of language background (F[1, 28] = 1.85, $MSE = 0.01$, $p = .184$, $\eta^2_p = .062$) or interaction between language background and trial type (F[1, 28] = 0.01, $MSE = 0.01$, $p = .934$, $\eta^2_p < .001$). Thus, we did not find evidence that infants noticed the speaker switching languages when using pupil dilation as a proxy for increased cognitive processing.

3.4 General Discussion

The current studies investigated whether bilingual and monolingual infants keep track of who speaks which language (i.e., spontaneously form a speaker-language association), by testing whether they would detect when speakers switched languages. In Study 1, we tested 5- and 12-month-olds using an auditory-only task and found no evidence that infants made an association between speaker and language spoken. In Study 2, we tested 12- and 18-month-olds and again found no evidence that infants made associations. In both studies, there was no difference between bilingual and monolingual infants. The ability to spontaneously form speaker-language associations relies on several prerequisite abilities, any of which could have contributed to infants' failure in the current study: (1) discriminating languages, (2) discriminating people, and

(3) associating a specific language with a specific person. We will discuss each of these in turn, before turning to the role of language experience and task modality.

First, infants must be able to discriminate between different languages in order to associate different languages with different people. Our study used English and French, which have different rhythmic and stress patterns, cues which are known to support language discrimination (Gasparini et al., 2021) Moreover, numerous studies across ages and paradigms have shown that infants can readily discriminate English and French, for example 2-month-olds in auditory language discrimination tasks (Dehaene-Lambertz & Houston, 1998), 4- to 8-month olds in visual language discrimination tasks (Weikum et al., 2007), as well as looking-while-listening tasks in older infants (e.g., Byers-Heinlein et al., 2017). Although we did not test infants' ability to discriminate our particular stimuli, based on these past studies, including some which used similar stimuli (Weikum et al., 2007), we expect that infants were able to discriminate between English and French in our study and that their failure at our task was not due to an issue with language discrimination.

Second, infants must be able to discriminate different individuals in order to associate them with a language. Our study used one male and one female speaker, given that male and female voices are discriminated early in development (Johnson et al., 2011; Miller, 1983). In Study 2, infants heard and saw the speakers, making it even less likely that they did not distinguish between speakers. We think it is therefore unlikely, but not impossible, that infants failed to discriminate between the two speakers.

Third, it could be that infants discriminated languages and the speakers, but that they struggled to make the association between a speaker and the language spoken. The ability to make associations is crucial across many aspects of learning, including learning about the sounds, words, and rules of a language (Saffran & Kirkham, 2018). However, infants make some types of associations more readily than others. For example, infants can learn that some speakers systematically pronounce some words differently, and generalize that to new situations (Weatherhead & White, 2016). Moreover, even within the same apparent task such as word-object associations, infants succeed in some paradigms (for review, see Tsui et al., 2019), but not in others (Gonzalez-Barrero et al., 2021). It may be that speaker-language associations are relatively difficult to form, or are better formed under other learning circumstances than the one we tested here. For example, infants might require a high degree of familiarity with speakers before they form a speaker-language association, and therefore might do so with well-known speakers in everyday life, but not spontaneously in a short laboratory task.

Our study also explored two potential moderators of infants' performance: the infants' language experience that they bring to the task (bilingual vs. monolingual), and the modality of the task. Our results did not show a difference for either of these two factors. With regards to language experience, we had expected to find that both bilingual and monolingual infants would form person-language associations, though possibly through different mechanisms. Bilingual infants are exposed to two languages regularly, and we therefore expected them to pay attention to language switches and possibly to language-person associations. We expected monolinguals, on the other hand, to notice when a speaker switches out of or into their native language. We, however, found that neither group appeared to learn the person-language associations, which again could be due to different underlying processes. Moreover, while our exploratory analyses provided some evidence that the direction of the language switch (into the dominant vs. the non-dominant language) affected performance, the direction of the effect was not consistent across ages and in the robustness analyses, and thus further research is warranted before interpreting this result.

With regards to modality, we expected infants to have an easier time learning the association in the audiovisual task, which provided more cues to the speaker's identity (the characteristics of both their voice and their face), and to the properties of the language (both in the auditory modality and information on the face and lips) than the auditory-only task. Indeed, it is hypothesized that redundant information from auditory and visual modalities is important for infants' development (Bahrick et al., 2004). For example, both auditory and visual cues contribute to performance in speech-in-noise tasks (Lalonde & Werner, 2019) and language discrimination (Sebastian Galles et al., 2012; Weikum et al., 2007). Further, a number of studies have shown that infants and children attend to the mouth of a speaker (e.g., Lewkowicz & Tift, 2012; Morin-Lessard et al., 2019), presumably to aid comprehension. Thus, the audiovisual condition provided infants with more information overall, and was potentially more engaging. However, we found no difference between the auditory-only and audiovisual studies, as infants did not learn the speaker-language association in either condition. This result is not unprecedented in the literature. In a study of language discrimination, Birulés et al. (2018) reported that bilingual 4-month-olds, when given audiovisual cues, did not discriminate Catalan and Spanish, a language pair that other studies have shown to be discriminable with auditory-only cues at the same age (Bosch & Sebastian Galles, 2001). Another study found that in some situations, seeing a speaker's mouth did not improve performance in a word recognition task (Singh et al., 2021). Together with the results of our study, this indicates that more information may not always result in better performance in infant learning (see also Newport, 1990).

Our research question also intersected with the rich literature on monolingual infants' use of language spoken as a social clue. At first glance, the lack of evidence for infants' ability to associate a person with a language seems contradictory with the findings that infants change their behaviour when observing people speak a certain language. However, there are several differences between our and their studies to consider. First, many of the previous studies tested preference for people speaking a certain language (e.g., Buttelmann et al., 2013; Kinzler et al., 2007) rather than infants learning of and recalling associations between a person and a language. In our task, infants had to learn the association between a speaker and a language, and then during the test trials notice that the speaker switched to a different language, a potentially demanding task. Second, other studies have relied on experiences that infants have made in their daily lives, such as about the connection between race and language (May et al., 2019; Weatherhead & White, 2018). For example, 3-year-olds make assumptions about the language individuals speak when shown pictures of people of different races, but not when shown pictures of old and young people (Hirschfeld & Gelman, 1997). This connection between race and language, learned from everyday experience, is likely stronger than associations learned in the lab. In our study, people of different genders spoke different languages, which may not be a typical regularity in the outside world. For regularities that are typically consistent, such as voice and visual cues to gender, these associations are learned in the first year of life (Hillairet de Boisferon et al., 2015; Walker-Andrews et al., 1991). We chose to use a male and a female speaker to maximize the differences between each speaker, however future studies could test for example the infants' own parents to tap into infants' knowledge of the language preferences of the people around them (see also Molnar & Davidson, 2017). These discrepancies between our studies' results and previous studies also highlight the importance of integrating findings across sub-disciplines of cognitive and social child development.

3.4.1 Implications and Future Directions

Our results have several implications for understanding bilingual language acquisition. First, a popular idea in bilingual parenting is that the one-person-one-language approach is the best way to support language acquisition (Barron-Hauwaert, 2004). One assumed benefit of this approach is that it helps infants to discriminate and separate their languages, as they associate different speakers with different languages. However, our results suggest that infants do not seem to readily track who speaks what language, although our study does not address whether infants might form such associations over the long term with their primary caregivers. Thus, the one-parent-one-language approach may not provide this purported benefit. Indeed, our results support the general finding that bilingual infants learn from a multitude of family language configurations providing they have sufficient high-quality experiences in each language. Second, it is possible that the way caregivers switch between languages already provides infants with helpful cues to detecting the switch. For example, studies of language mixing in parents show that parents mostly switch languages between sentences rather than within sentences (Kremin et al., in press), which may help infants' comprehension (Byers-Heinlein et al., 2017). Given that we did not find statistically significant results in our task, there is an open opportunity for future research. One avenue would be to use a different paradigm that does not rely on infants being *surprised* that a person is switching languages, as this may be a common occurrence in infants' lives. Instead, sensitivity to language switching could be tested by introducing two speakers who each speak a different language, as in our study, and then showing pictures of both speakers while infants hear one language. If infants look more at the person previously associated with the language they are hearing, this would indicate that they made the association between the language and the speaker.

Furthermore, it might be that the differences between languages in certain tasks are less salient than we would assume based on language discrimination studies. A recent study found that infants did not discriminate between English and French when hearing single words that were either in a single language (dog… milk) or alternated between languages (dog... lait [fr. milk] Schott et al., 2021). Whereas in a looking while listening task, 20-months-olds did notice when the language switched between English and French (Byers-Heinlein et al., 2017). The study on single-word language discrimination used an infant-directed speech comprehension task, whereas our study was modelled after language discrimination studies and therefore used adult-directed stimuli that were not designed to be understood by infants. It is possible that language discrimination happens as a "by-product" of higher-level processes like lexical access of words that are organized by language, which we would not have tapped into using our task, as our task was not designed to require lexical access. Our results suggest that language discrimination is more nuanced than previously thought based on early language discrimination studies.

3.4.2 Conclusion

We showed in two studies that infants do not appear to spontaneously form associations between a person and the language they speak. This finding hints at the potentially long developmental path for bilingual infants to move from simple discrimination of their languages to more complex forms of language discrimination, separation, and social use (Byers-Heinlein, 2014). This work highlights a need to integrate perceptual and social approaches, to uncover the mechanisms at play in early bilingual language learning.

4 Bilingual and Monolingual Toddlers have Similar Sensitivity to Mispronunciations in Cognate and Non-Cognate Words

4.1 Introduction

Many children are exposed to two languages from birth, and have to learn two sets of sounds and two sets of words, one in each language. Words across bilinguals' languages can vary in their relationship: some cross-language synonyms sound very distinct (e.g., English apple and French *pomme*) while others sound similar (e.g., English chocolate and French *chocolat*). Cross-language synonyms (also called translation equivalents) that share both meaning and overlapping sounds are called cognates. The current study asked whether the substantial phonological overlap might cause bilinguals to represent cognates differently, for example in less phonological detail than non-cognates, based on prior research indicating that this may be the case (Ramon-Casas et al., 2009; Ramon-Casas & Bosch, 2010). We investigated bilingual toddlers' comprehension of correctly pronounced and mispronounced cognates and non-cognates in a looking-while-listening task. We also tested monolingual toddlers, who served as a comparison group.

4.1.1 Coping with Variability in Word Learning and Phonological Encoding

When recognizing words, both bilingual and monolingual children have to navigate which sound changes are meaningful and which are not. For example, some phonological changes that do not affect word meaning should be accommodated, such as those resulting from different speakers, intonation, and accents. At the same time, other phonological changes should be distinguished, such as the vowel difference between "ball" and "bowl". To test infants' phonological representations, researchers often use a mispronunciation paradigm where a word is either pronounced correctly or incorrectly (e.g., dog or *bog). The duration of infants' looking to the target object when it is labelled correctly or incorrectly is then compared. If infants look to the target object less when hearing the incorrect pronunciation, this indicates that they noticed the mispronunciation (Swingley, 2005), and thus their phonological representation did encode such details. Monolingual infants have been shown to be sensitive to mispronunciations in many studies (Von Holzen & Bergmann, 2021).

There is mixed evidence as to whether navigating two sets of speech sounds impacts bilinguals' encoding of phonological detail in words. Some studies of mispronunciation detection show that bilinguals, like monolinguals, are sensitive to mispronunciations (Singh et al., 2020; Wewalaarachchi et al., 2017), while others find that bilinguals are less sensitive to mispronunciations (Ramon-Casas et al., 2009). There is prior evidence that bilingual and monolingual infants differ in whether they can use subtle phonological differences in word learning tasks (Fennell et al., 2007), where bilinguals succeed at a later age than monolingual infants. Bilingual infants are able to use subtle phonological differences to guide novel word learning only if the speaker matches the infant's language environment (Fennell & Byers-Heinlein, 2014; Mattock et al., 2010). Overall, this body of work suggests that both bilingual and monolingual infants are sensitive to phonological detail, but that this sensitivity is more dependent on the characteristics of the task in bilinguals compared to monolinguals (see also Curtin et al., 2011).

4.1.2 Processing Cognates versus Non-Cognate Translation Equivalents

Bilingual word learning is unique in that children need to learn two words for every concept – one in each of their languages. Cognates are a subset of translation equivalents that share an etymological origin, and thus sound and/or are spelled similarly (e.g., chocolate/*chocolat*). As our focus is on pre-literate infants, we concentrate here on sound overlap, rather than spelling. Cognates exist on a dimension from (almost) identical to vaguely related. Some cognates sound very similar (carrot - *carotte*), while others sound less similar (brush - *brosse*). Similarly, languages vary in how many cognates they share: some languages share many cognates (e.g., Spanish and Catalan), and others share fewer cognates (e.g., English and Japanese).

Many studies of adults show that cognates are read faster than non-cognates (e.g., Dijkstra et al., 2010). This is thought to occur because orthographic (and phonemic) features of both languages are activated when reading cognates, and this increased activation facilitates processing (Dijkstra & van Heuven, 2002). Particularly relevant here are studies where the cognate effect cannot be explained by purely orthographic similarity. Cognates exist for languages with different scripts such as English and Japanese, e.g., Japanese: レモン /remoN/, English: lemon /ˈlɛmən/). For bilingual adults who speak different-script languages, cognate effects cannot be explained based on orthographic similarity. And indeed, cognate effects can be found even when testing different-script bilinguals (Gollan et al., 1997; Nakayama et al., 2013; Voga & Grainger, 2007; but see Kim & Davis, 2003). In summary, studies on adults' processing of cognates have led to the creation and refinement of models of bilingual vocabulary organization (e.g., Dijkstra & van Heuven, 2002; Miwa et al., 2014).

While much less studied than adult cognate processing, there are some studies of cognate processing in children. A study of school-aged children shows cognate words are read faster than non-cognate words (Bosma & Nota, 2020). However, of particular interest for our task are studies of processing differences for cognates versus non-cognates in pre-literate children. Toddlers are reported to produce more cognates compared to non-cognates, suggesting that phonological overlap supports word learning (Bosch & Ramon-Casas, 2014; Mitchell et al., 2022). More closely related to word recognition, a looking time study of German toddlers learning English showed that English but not German word recognition was modulated by cross-language similarity (Von Holzen et al., 2018). However, the toddlers tested in this study learned English after they had learned German for at least a year, and there is evidence from adults that cognates facilitate second language learning (e.g., De Groot & Keijzer, 2000; Rogers et al., 2015). Thus, the results from this study may not apply for bilingual children learning two languages at the same time. A study of Spanish-English preschoolers found different patterns for English-dominant, Spanish-dominant, and balanced bilinguals, where the children dominant in one language showed a difference in pointing accuracy to cognate and non-cognate words. This difference was however not found for balanced bilinguals (Pérez et al., 2010). Thus, there are few studies testing whether preliterate children process cognates differently from non-cognates, and the work that does exist is contradictory as to whether cognates and non-cognates are processed differently.

4.1.3 Phonological Detail in Cognate Representations

Another unanswered question is whether bilinguals' word representations are affected by the properties of the word pairs to be learned. We have some evidence that the properties of the words to be learned matter in word learning studies in the lab. Bilinguals can learn labels for two novel objects when they sound different (e.g., *lif-neem*) at 14 months (Byers-Heinlein et al., 2013), but only show learning for similar-sounding words (e.g. kem/gem) at 17-20 months (Fennell & Byers-Heinlein, 2014; Mattock et al., 2010; Singh et al., 2018). Presumably, bilinguals can differentiate the sounds in the similar-sounding pair (e.g., k/g distinction) at 14 months, but they do no not use their ability to differentiate these similar-sounding words in the word learning task until 17 months. However, when instead of using consonant distinctions, vowels are varied, bilingual infants can show better word learning than monolinguals (Singh et al., 2018). Thus, sometimes bilingual and monolingual infants use subtle sound distinctions to different degrees in word learning tasks. The present research will ask whether a similar process happens in bilingual infants for similar and different-sounding word pairs that exist naturally, that is, for cognate words.

The small phonological variations between the cognate labels may lead bilingual infants to initially only encode the shared sounds in detail and underspecify the non-shared sounds, leading to a less detailed word representation for cognates compared to non-cognates (see also Ramon-Casas et al., 2017). Several studies have examined whether bilingual infants represent cognates with the same phonological detail as non-cognates. A series of studies on Catalan-Spanish bilingual toddlers found that for cognate words, 18- to 24 month-olds did not differentiate between correctly pronounced and mispronounced cognate words that varied in vowel distinction only phonemic in one of the bilinguals' languages (Ramon-Casas et al., 2009). For non-cognate words, bilingual children did differentiate between correctly pronounced and mispronounced words (Ramon-Casas & Bosch, 2010). This suggests that cognates were encoded with less phonological detail. It is important to test whether the finding that bilinguals' representations of cognates are underspecified generalizes to other populations, or whether this result is specific to infants learning highly similar languages such as Spanish and Catalan.

4.1.4 The Current Study

We tested whether bilingual toddlers process cognates differently from non-cognates, and whether their representation of phonological detail is affected by cognate status. This has not been tested in bilingual toddlers that learned both of their languages from birth, and who learn languages that are less similar than Spanish and Catalan, for example. We tested English-French bilinguals. French and English are from different language families, thus these languages' sound systems are different, yet centuries of contact between the two languages has led to a fair number of shared cognate words (Schepens et al., 2012). Yet, English and French have much fewer cognates than for example Spanish and Catalan, which makes English and French a good test case for cognate effects in less closely related languages. We chose sets of cognate and non-cognate nouns which were matched on word length and age of acquisition. Based on the mean age of acquisition of those words, we decided to test 24- to 30-month-olds. We used an eye-tracking looking-while-listening paradigm, where toddlers see two objects on the screen and hear one labelled, and we measure how long toddlers look at the target word compared to the distractor. A mispronunciation paradigm was used to probe toddler's representations of phonological detail, and thus on half of the trials, toddlers heard the target word mispronounced, and their looking time was compared to their performance on correctly pronounced trials. In this paradigm, toddlers with a more fine-grained encoding of phonological detail should perform worse when hearing mispronounced compared to correctly pronounced words. We used both vowel and consonant mispronunciations. While several studies postulate that consonants affect word recognition more than vowels (e.g., Hochmann et al., 2011; Poltrock & Nazzi, 2015), cognates often vary across languages mostly in their vowels, making it important to test vowel mispronunciations as well. We furthermore measured pupil dilation in response to hearing mispronunciations, as an indicator of increased cognitive processing (Sirois & Brisson, 2014; Tamási et al., 2017).

We tested three predictions with our study. The first prediction was that bilingual toddlers would be less sensitive to mispronunciations than monolingual toddlers, motivated by prior research on Catalan-Spanish toddlers (Ramon-Casas et al., 2009). We tested this by comparing looking times to the target object when the target was being correctly pronounced versus when it was mispronounced. We also measured pupil dilation responses to hearing each trial type. We expected an increase in pupil size for mispronunciations compared to correct pronunciations, which should be more pronounced for bilingual compared to monolingual toddlers. The second prediction was that bilingual toddlers would show better word recognition for cognate compared to non-cognate words, as evidenced by longer looking time to the target word on cognate compared to non-cognate trials. This prediction was motivated by adult literature showing cognate effects, as well as limited evidence from young children (Squires et al., 2020; Von Holzen et al., 2018). Monolingual toddlers serve as a control group here, as we should not see differences between these two word types for this group. Longer looking times to the target for correctly pronounced relative to mispronounced trials would indicate more detailed phonological encoding. The third prediction was that bilingual toddlers would be less sensitive to mispronunciations for cognate words compared to non-cognate words, as suggested for Catalan-Spanish toddlers (Ramon-Casas et al., 2009; Ramon-Casas & Bosch, 2010). Again, we predicted no differences between mispronunciations for cognate and non-cognate words for monolingual toddlers.

The preregistration for the current study can be found at https://osf.io/u4eqc.

4.2 Method

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Human Research Ethics Board of Concordia University. Parents were informed of the study's purpose and procedure, and provided written consent prior to participation (see Appendix A).

4.2.1 Participants

The final sample consisted of 24 bilingual and 27 monolingual toddlers aged between 24 and 31 months (see also Table 4.1). Participants were recruited from government birth lists as well as a database of interested parents, and were tested between June 2018 and June 2019. Our planned sample size was 24 participants per language group, which was chosen because it seemed a feasible number of bilingual participants to test. For both groups, we ended up testing more than 24 participants due to a delay between participant recruitment and subsequent checking which participants met inclusion criteria. We included all participants in the final analysis who met the pre-registered inclusion criteria.

Toddlers were assigned to language groups using the Language Exposure questionnaire (Bosch & Sebastian Galles, 2001) using MAPLE (Byers-Heinlein et al., 2020) to elicit information about current and lifetime language exposure from parents (see Appendix B). From this information, we calculated an overall % of exposure to English and French as well as all other languages in the toddlers' life. To be classified as bilingual, participants had to be exposed to at least 25% English and 25% French, and less than 10% of an additional language. Bilingual participants in our sample were exposed to their dominant (most heard) language at an average of 63% (range: 52-75%), and their non-dominant language at an average of 36 % (range: 25-48%). The dominant language was English for 50 % of bilingual participants. To be classified as monolingual, toddlers had to be exposed to English at least 90% of the time, and on average were exposed to English at 97% (range: 90-100%). Participants were reported to be 59% of European descent, 18% multiple responses, 10% Caribbean, all other categories made up less than 10% of the total participants. Maternal education was on average 16.0 years (SD = 2.4).

To get to this final sample, we tested 21 additional participants who did not meet our pre-registered inclusion criteria. We excluded participants who were born prematurely (<37 weeks) or with low birth weight (< 2500g, N = 3), and those who did not fit our language criteria (N = 18). No participants were excluded for major health issues that may affect their hearing/vision or typical development. For all participants who met the general criteria for inclusion, we verified that they contributed enough eye tracking data to be included, and excluded four participants (3 bilingual, 1 monolingual) who contributed less than 2 trials of each of the conditions (correctly pronounced cognates, mispronounced cognates, etc.).

Table 4.1 Sample Characteristics by Language Group

	Bilingual	Monolingual
n	24	27
Mean age in months (*SD*)	27.08 (2.15)	26.25 (2.1)
Age range in months	24.1-31.4	24.13-31.4
% Girls	0.38	0.52

4.2.2 Auditory and Visual Stimuli

Auditory and visual stimuli can be found on the Open Science Framework (https://osf.io/n9uv4/?view_only=82e1bb8b5b204e15814bb8fc0801995b). We chose six English words that had cognate translation equivalents and six English words that had non-cognate translation equivalents. The cognate and non-cognate words can be seen in Table 4.2. Each word was paired with another to form a target-distractor pair, and over the course of the study each word appeared equally often as a target and as a distractor. Cognate and non-cognate words lists were matched on their typical age of acquisition by using archival data from our lab as well as CDI norms (Jørgensen et al., 2010) and were matched on syllable and phoneme length.

Table 4.2 English Stimuli, Mispronunciation Stimuli, and French Translations

	English stimulus	Mispronunciation stimulus	French translation
Cognates	banana /bəˈnæ.næ/	boonana /buˈnæ.næ/	banane /ba.nan/
	bowl /boʊl/	coal* /koʊl/	bol /bɔl/
	chocolate /ˈt͡ʃɔklət/	chucklat /ˈt͡ʃuklət/	chocolat /ʃɔ.kɔ.la/
	giraffe /d͡ʒəˈɹæf/	viraffe /vəˈɹæf/	girafe /ʒi.ʁaf/
	table /ˈteɪbəl/	tebel /ˈtɛbəl/	table /tabl/
	pizza /ˈpitsə/	kitza /ˈkitsə/	pizza /pi.dza/
Non-Cognates	butterfly /ˈbʌɾɚflaɪ/	bitterfly /ˈbɪɾɚflaɪ/	papillon /ˈpæpɪjɔ̃/
	cookie /ˈkʊki/	khecky /ˈkɛki/	biscuit /biskɥi/
	foot [fʊʔ]	soot* [sʊʔ]	pied /pje/
	monkey /ˈmʌŋki/	vonkey /ˈvʌŋki/	singe /sɛ̃ʒ/
	mouth /maʊθ/	gouth /gaʊθ/	bouche /buʃ/
	window /ˈwɪndoʊ/	wendow /ˈwɛndoʊ/	fenêtre /fnaɛ̯t/

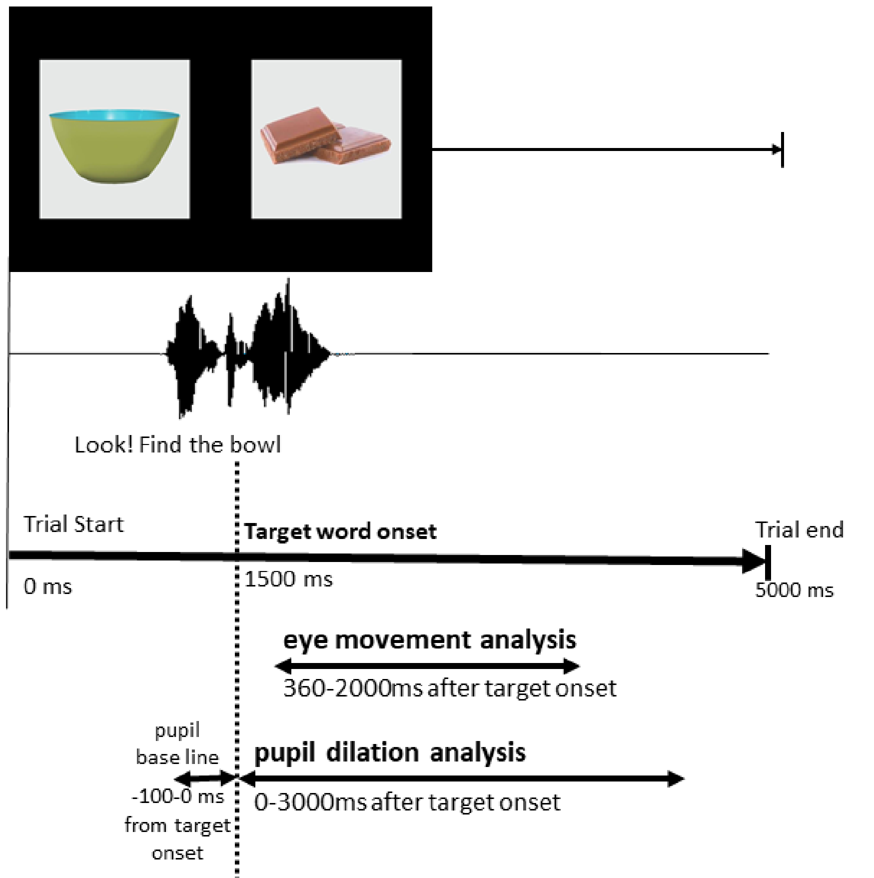
Note. French translations are included for reference only, all stimuli in the study were presented in English. IPA transcriptions added below each word. Words marked with an asterisk are real words in English, but are unlikely to be known by toddlers of this age.

Mispronunciations were created by changing a sound in the first syllable of each word (see Table 4.2). Mispronunciations were one or two feature changes, and the size of the mispronunciation was not significantly different for cognates ($M = 1.67$) and non-cognates ($M = 1.83$, $t\ (9.5) = -0.62$, $p = 0.550$). When crafting mispronunciations, the word onset could not be changed to be similar to the onset of the distractor in either English or French, e.g., if the child sees a bowl and chocolate, the mispronunciation of bowl would not be /ˈt͡ʃoʊl/, that is, using the first sound in chocolate. Half of the mispronunciations affected the onset consonant (e.g., giraffe changed to viraffe), and half affected the first vowel after the onset consonant (e.g., banana changed to boonana). Both consonant and vowel mispronunciations were used, as there is some debate in the literature as to which is more disruptive to word recognition (Delle Luche et al., 2014; Floccia et al., 2014; Poltrock & Nazzi, 2015; Singh et al., 2018). No target words had a vowel onset. The mispronunciations were constructed such that the mispronounced word does not form a noun competitor likely to be known to children at this age (i.e., bowl would not be changed to ball). However, due to the other constraints on forming mispronunciations, two mispronunciations ended up sounding like existing words (foot - soot and bowl - coal) that are unlikely to be known to toddlers. We tested all children in English, instead of creating English and French stimuli, because it is difficult to create mispronunciations of comparable magnitude in English and French, so this choice provided increased experimental control.

Stimuli were recorded by a fluent English-French bilingual female speaker speaking in a Montréal native accent in both English and French. Object labels were recorded in English, both with correct pronunciation (e.g., banana) and with a mispronunciation in the first syllable (e.g., boonana). Care was taken when recording that the mispronunciation only affected one sound. All stimuli were recorded with the same sentence frame "Look! Find the …!", and no splicing or editing was applied.

High-resolution photographs of objects corresponding to cognate and non-cognate items against a light grey background (680×720 px) were used to present during the study. On each trial, two objects were displayed side-by-side on a black background (see Figure 4.1). Thus, each item was used as both target and distractor across the whole study. Cognates were paired with cognates, and non-cognates were paired with non-cognates. The full set of stimuli can be seen in Table 4.2. We chose pairings so that for non-cognate words, the onset and rhyme of both the English and French labels were as different as possible. Images that were displayed together were processed using a custom Matlab script to equate for luminance, to improve the interpretability of pupil dilation results.

Figure 4.1 Timeline of Trial Structure



4.2.3 Procedure

Participants were tested in a dimly lit sound-attenuated room. Infants sat on their parent's lap facing the screen of a Tobii T60-XL eye tracker (Tobii® Technology, 2010) that gathered eye-gaze data. Parents wore darkened sunglasses and listened to music on headphones to minimize their interaction with the infants during the study. The experimenter controlled the study from an adjacent room using Tobii Studio software and was unaware of which object was named on each trial. Before stimuli presentation, the eye-tracker was calibrated to the participants' eyes using a five-point infant calibration routine. A bouncing circle changing in colour and shape was presented before each trial to redirect the infant's gaze to the center of the screen. Infants saw 24 trials. On each trial, toddlers heard a speaker name one of the objects (the target) on the screen: "Look! Find the chocolate/bowl!". Each object served as the target for half of trials (12), and as distractor for the other half of trials.

Mispronounced and correctly pronounced trials as well as cognate and non-cognate naming trails were pseudorandomized such that no more than two trials of the same type appear consecutively. The location of the target was also counterbalanced such that targets appeared equally on the left and right sides.

After completing the eye tracking portion of the study, parents completed questionnaires on their language mixing practices (Appendix C) and demographic information (Appendix D) as part of standard lab practice. Parents were then debriefed on the study, and their child received a certificate and a small toy as a thank you.

4.3 Results

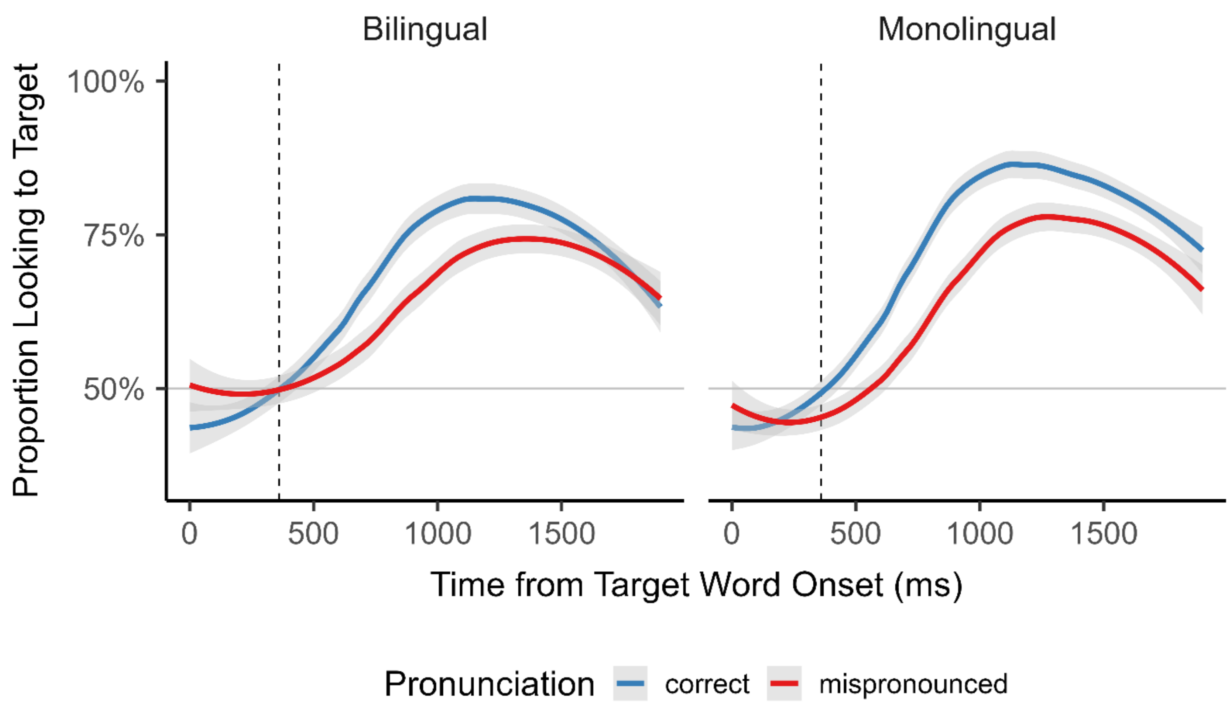
4.3.1 Looking Time: Analytic Approach

The area of interest for the current study was the 680×720 px area around the target object. The dependent variable was toddlers' looking to the area of interest during the pre-registered analysis window of 360-2000 ms after onset of the target word divided by their total looking time to the target and distractor. As described in the participants section, we removed four participants who did not contribute at least two trials of each of the four trial types (correctly pronounced cognate, mispronounced non-cognate, correctly pronounced non-cognate, mispronounced non-cognate). After exclusion, the remaining 51 participants contributed an average of 21.4 out of 24 trials. We then compared toddlers' looking to the target to chance, an analysis that was inadvertently omitted from our pre-registration, but that is commonplace analysis in word recognition studies (Von Holzen & Bergmann, 2021). Our main pre-registered analysis was an ANOVA with cognate status and mispronunciation as within-subject variables and language background as a between-subjects variable. The time course of looking to the target is shown in Figure 4.2.

4.3.2 Comparison to Chance

We compared toddlers' looking to the target object to chance (50% in our two-image display) using two-sided one-sample t-tests. We found that both bilinguals and monolinguals looked at the target word above chance, across both correctly pronounced and mispronounced trials, as well as across cognate and non-cognate trials (all *p*s < 0.001, see Table 4.3). This shows that toddlers were able to recognize the target words, and that mispronunciations did not completely derail children's comprehension.

Figure 4.2 Looking to Target for Mispronounced and Correctly Pronounced Trials, by Group



Note. Dashed line indicates onset of window of analysis (360 ms after target word onset). Proportion of looking to target is smoothed using a loess fit. The grey area indicates 95 % CI.

4.3.3 Looking Time ANOVA

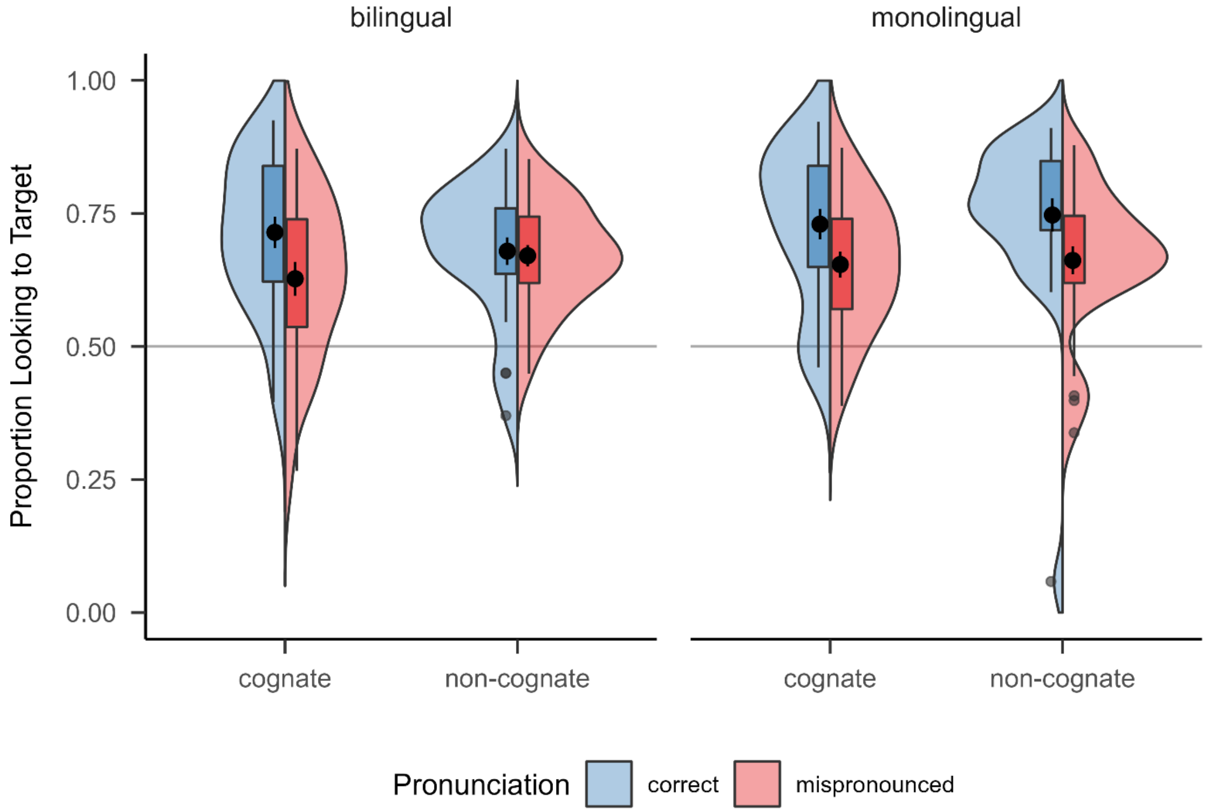
The proportion looking to target averaged across the window of analysis can be seen in Figure 4.3. We calculated a 2 × 2 × 2 ANOVA with cognate status (cognate, non-cognate) and mispronunciation (correctly pronounced, mispronounced) as within-subject variables, and language background (monolingual, bilingual) as a between-subjects variable. The results can be seen in Table 4.4. We found a main effect of mispronunciation, with less looking at the target object for mispronounced (M = 65%) compared to correctly pronounced words (M = 72%). There was no main effect of cognate status nor any interaction between cognate status and mispronunciation, indicating that there was no difference in word recognition between cognate and non-cognate nouns, whether correctly pronounced or mispronounced. There was no evidence that language background was related to children's performance, as all main effects and interactions that involved language group were non-significant.

Table 4.3 Looking Time to Target Compared to Chance, by Group

	Bilinguals			Monolinguals		
Pronunciation	M	Cohen's *d*	*t*-test	M	Cohen's *d*	*t*-test
			Cognates			
correct	0.71	1.49	t (23) = 7.32, p < 0.001	0.73	1.55	t (26) = 8.06, p < 0.001
mispronounced	0.63	0.82	t (23) = 4.00, p < 0.001	0.65	1.22	t (26) = 6.32, p < 0.001
			Non-cognates			
correct	0.68	1.40	t (23) = 6.88, p < 0.001	0.75	1.51	t (26) = 7.87, p < 0.001
mispronounced	0.67	1.73	t (23) = 8.46, p < 0.001	0.66	1.19	t (26) = 6.18, p < 0.001

To compare mispronunciation sensitivity for cognates and non-cognates for bilingual children only, we calculated a 2 × 2 ANOVA with cognate status and mispronunciation as within-subject variables. The main effect of mispronunciation was trending towards but did not reach statistical significance ($F[1, 23] = 4.16$, $MSE = 0.01$, $p = .053$, $\eta^2_p = .153$). There was no main effect of cognate status ($F[1, 23] = 0.03$, $MSE = 0.01$, $p = .863$, $\eta^2_p = .001$), indicating that bilingual toddlers did not look more to the target when they heard a cognate compared to a non-cognate noun. We found no interaction between cognate status and mispronunciation ($F[1, 23] = 3.04$, $MSE = 0.01$, $p = .095$, $\eta^2_p = .117$), indicating that cognate status did not affect bilinguals' sensitivity to mispronunciations.

Figure 4.3 Proportion of Looking to Target for Analysis Time Window



Note. Black dots with whiskers inside the boxplots show the mean and the standard error of the mean. Violin plots show the distribution of values.

Table 4.4 ANOVA on Proportion of Looking to Target

Predictor	df_{Num}	df_{Den}	F	p	η^2_g
Mispronunciation	1	49	14.44	< .001	.05
Cognate Status	1	49	0.26	.611	< .01
Language Background	1	49	1.01	.320	.01
Cognate Status × Mispronunciation	1	49	0.80	.377	< .01
Language Background × Mispronunciation	1	49	0.91	.346	< .01
Language Background × Cognate Status	1	49	0.07	.796	< .01
Language Background × Cognate Status × Mispronunciation	1	49	1.52	.223	.01

Note. df_{Num} indicates degrees of freedom numerator. df_{Den} indicates degrees of freedom denominator. η^2_g indicates generalized eta-squared.

4.3.4 Pupil Dilation ANOVA

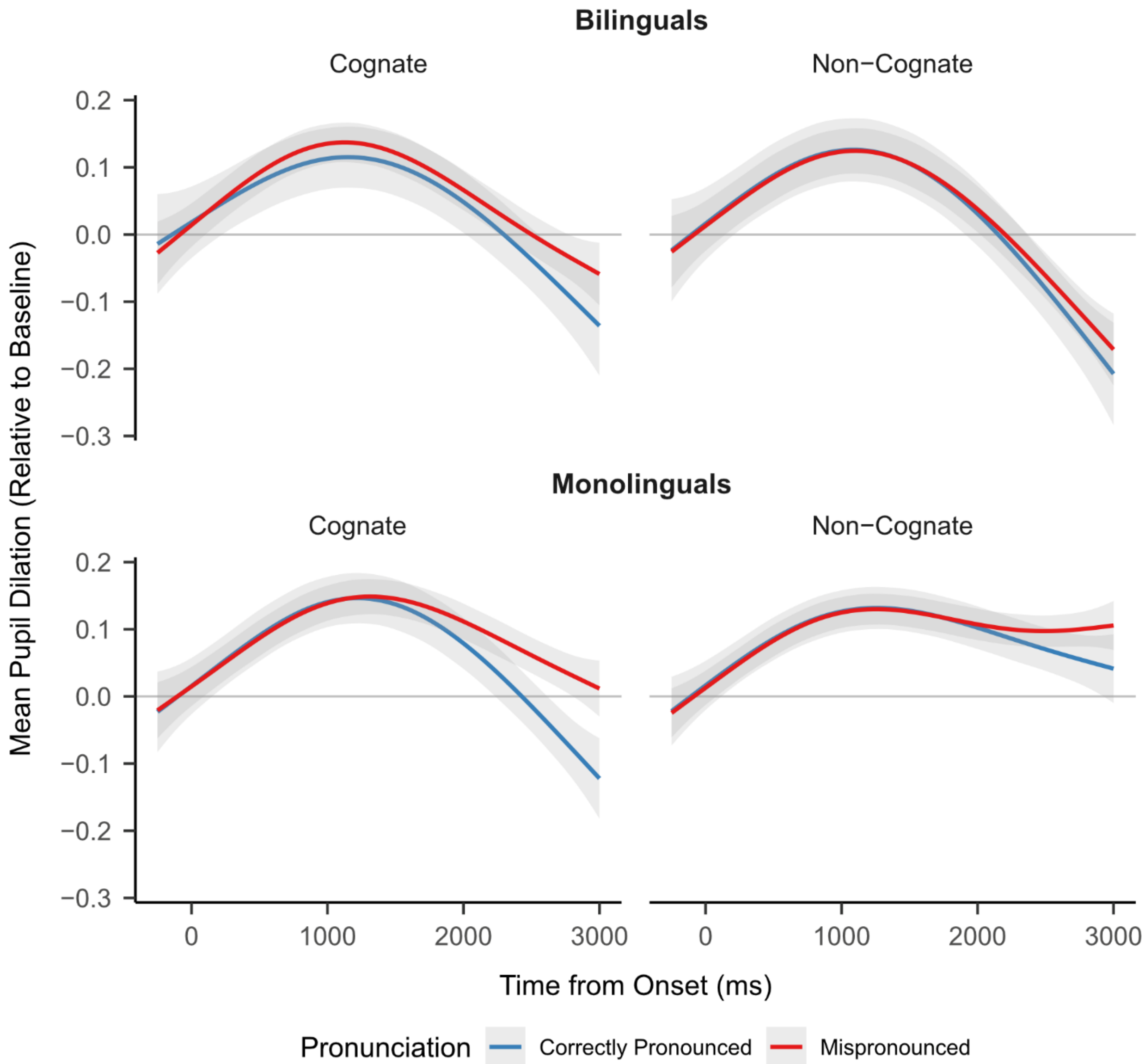
To analyze change in pupil size, we used the pupillometryR package (Forbes, 2020) and followed the steps recommended for pupillometry (Sirois, 2018). The window of analysis was pre-registered to start at the onset of the target word and end 3000ms after. As the correlation between the left and right pupil size was high (r = .95), in cases where pupil size for one eye only was available, we regressed the size of the missing pupil from the available side. We then calculated the average pupil side across both eyes, and downsampled the data to 250 ms time bins to smooth and make the size of the dataset more manageable. We removed trials with less than 750ms of usable data, which removed 29 trials or 2.6% of trials. Following this, we removed all participants who had less than 2 usable trials on each of the four trial types, which resulted in removing 6 participants (10.9% of participants). This resulted in a final sample for pupillometry analyses of 24 bilinguals and 25 monolingual toddlers. We then filtered data to remove artifacts like implausibly fast pupil dilations using a Hanning filter. Data was then interpolated across blinks and other periods of missing data using a cubic interpolation. Finally, data were baseline-corrected using a baseline of 100 ms before the onset of the target word, to remove variability in pupil sizes between participants and trials.

Change in pupil size can be seen in Figure 4.4. A 2 × 2 × 2 ANOVA with cognate status (cognate, non-cognate) and mispronunciation (correctly pronounced, mispronounced) as within-subject variables, and language background (monolingual, bilingual) as a between-subjects variable. The results can be seen in Table 4.5. Unlike for the looking time analysis, we found no main effect of mispronunciation (p = .137, η^2_p = .046), indicating that toddlers' pupil size did not differ between correct and mispronounced trials. For the remaining main effects and interaction, we also found no statistically significant main effects or interactions (all ps > .131, see Table 4.5). This suggests that pupil dilation was not associated with the mispronunciation or cognate status manipulation.

We also conducted a separate ANOVA for bilingual toddlers only, to test whether there are effects present in bilinguals but not in monolinguals that the three-way ANOVA did not have the power to detect in a three-way interaction. We again found no statistically significant effect of mispronunciation (F[1, 23] = 0.73, MSE = 0.01, p = .402, η^2_p = .031), or cognate status (F[1, 23] = 0.44, MSE = 0.02, p = .513, η^2_p = .019), or interaction between the two (F[1, 23] = 0.19,

MSE = 0.01, p = .665, η^2_p = .008). This suggests that there were no effects of mispronunciation or cognate status on pupil size present in bilinguals specifically.

Figure 4.4 Pupil Dilation for Mispronounced and Correct Trials, by Group and Cognate Status



Note. Mean pupil dilation is smoothed using a loess fit. The grey area around the smoothed line indicates 95 % CI.

Table 4.5 ANOVA with Change in Pupil Size as Dependent Variable

Predictor	df_{Num}	df_{Den}	F	p	η^2_g
Mispronunciation	1	47	2.28	.137	.01
Cognate Status	1	47	< 0.01	.964	< .01
Language Group	1	47	2.36	.131	.03
Mispronunciation × Cognate Status	1	47	0.53	.471	< .01
Language Group × Mispronunciation	1	47	0.02	.883	<.01
Language Group × Cognate Status	1	47	1.15	.289	<.01
Language Group × Mispronunciation × Cognate Status	1	47	0.01	.922	<.01

Note. df_{Num} indicates degrees of freedom numerator. df_{Den} indicates degrees of freedom denominator. η^2_g indicates generalized eta-squared.

4.4 Discussion

In the current study, we investigated bilingual toddlers' representation of phonological detail in word recognition and how it is impacted by cross-language similarity, specifically whether a word is a cognate (e.g., chocolate - *chocolat* [fr. chocolate]) or non-cognate (e.g. apple - *pomme* [fr. apple]). We tested 24- to 31-month-old bilingual toddlers in a preferential looking eye tracking task, and tested monolingual toddlers as a comparison group. We found no difference between bilingual and monolingual toddlers' representation of phonological detail, as both groups looked less to the target word when it was mispronounced, indicating that they noticed the change in sounds. Both groups nonetheless still looked above chance to the target when it was mispronounced, indicating resilience in children's word comprehension. Moreover, we found no differences in toddlers' recognition of cognate and non-cognate words, whether correctly pronounced or mispronounced.

4.4.1 Bilinguals and Monolinguals Sensitive to Mispronunciations

We had three predictions for our study. The first was that bilingual toddlers would be less sensitive to mispronunciations than monolinguals. This was based on previous research showing that bilinguals are less sensitive to certain vowel mispronunciations than monolinguals (Ramon-Casas et al., 2009). We instead found that both groups were sensitive to mispronunciations, and to a similar degree. Some other studies comparing bilingual and monolingual 24- to 27-month-olds have also reported that Singaporean toddlers learning English and/or Mandarin are sensitive to mispronunciations (Singh et al., 2020; Wewalaarachchi et al., 2017). However, these studies reported that toddlers did not look to the target above chance when the label was mispronounced, whereas we found that toddlers looked to the target above chance even in the mispronunciation condition. Studies with monolingual toddlers also mostly report above-chance looking even when hearing a mispronunciation (Moore & Bergelson, 2021; Von Holzen & Bergmann, 2021; Weatherhead & White, 2018), which is what we observed in our study. It is unclear why some studies show that hearing a mispronunciation completely disrupts word recognition and others do not. While some differences exist between those studies, such as the dependent variable or the use of unfamiliar objects as distractors, those were not found to affect mispronunciation sensitivity in a meta-analysis (Von Holzen & Bergmann, 2021). The size of the mispronunciation seems comparable between the studies, making it unlikely this was the cause of the difference observed. Other possible explanations are cultural differences between Western and Southeast Asian cultures (although see Moore & Bergelson, 2021 for a condition where North American toddlers were at chance on mispronunciations), or other subtle differences in experimental design not explored in the meta-analysis. Further investigation is needed to determine if this is a meaningful difference. Overall, we found no evidence for a difference between bilingual and monolingual toddlers in their sensitivity to mispronunciations.

We also investigated pupil size changes in response to mispronunciations, as an index of cognitive processing. Toddlers showed a similar pattern of pupil responses whether words were correctly pronounced or mispronounced. This contrasts with two studies of monolingual 30-month-olds which reported greater pupil dilation on mispronunciation trials, although these studies used a single fixation design whereas our study used a preferential looking design (Fritzsche & Höhle, 2015; Tamási et al., 2017). One study that used preferential looking with 24-month-olds reported mixed results, with a pupil dilation difference for bilinguals but not monolinguals (Tamási et al., 2016) It is possible that an effect of pupillometry is easier to observe in a single-fixation design, as pupillometry is sensitive to changes in luminance that could stem from switching fixation in preferential looking designs. This is supported by another preferential looking study of monolingual 30-month-olds (Tamási et al., 2019) that showed no difference between correctly pronounced words and mispronounced words with 1 or 2 phonetic feature changes, which matches with the results of our study. Overall, these findings suggest that, at least in preferential looking tasks, looking time might be more sensitive than pupil dilation in measuring toddlers' perception of mispronunciations.

4.4.2 No Difference Between Cognate and Non-Cognate Word Recognition

Our second prediction was that bilingual infants would be better at recognizing cognate words compared to non-cognate words, as we expected that sound overlap between the two forms of a cognate might boost the activation of the labelled concept. Instead, we found no difference between toddlers' recognition of cognate and non-cognate words. This contrasts with the results of a word recognition study of 18- to 53-month-olds who learned English as a second language starting at 12 months (Von Holzen et al., 2018). In this study, time course analyses showed that English cognates were recognized faster than English non-cognates, but cognates and non-cognates were recognized with similar efficiency in participants' dominant language, i.e., German. Thus, this hints that cognate effects in word recognition may depend on the nature of children's bilingual exposure. The revised hierarchical model (Kroll et al., 2010) can explain how cognate effects can occur differently in different groups of bilinguals. For unbalanced and/or sequential bilinguals, who have one strong language and a second weaker one, lexical access in the second language would go through the first language. Thus, for cognate words, the more efficient dominant-language representations would be activated alongside the non-dominant representation, effectively fast-tracking word recognition in the non-dominant language. By contrast, for balanced bilinguals, word representations in each language are equally strong and thus it is less likely that the form similarity of translation equivalents affects word recognition. This would explain why in our population of simultaneous and relatively balanced bilinguals we don't find an effect of cognate status, while Von Holzen et al. (2018) found a cognate effect in unbalanced sequential bilinguals only in the non-dominant language.

The absence of a cognate effect is somewhat surprising in the context that many studies have found this effect across a range of tasks in adults (e.g., lexical decision, reading, picture naming) and some evidence in tasks with children (e.g., picture-word identification, reading). While most of the studies testing cognate effects in adults are testing reading, there are a few studies also testing other modalities. For example, a cognate facilitation has been found in production (Muylle et al., 2022). In a study testing adults using a looking-while-listening task very similar to ours, bilingual adults looked more at the target object when the target was a cognate compared to a non-cognate (Andras et al., 2022). Thus, this suggests that testing auditory vs. visual word recognition cannot solely account for the discrepancy between our research and prior studies.

Indeed, our research group has previously found a cognate advantage in this same population of English-French bilingual toddlers we tested when measuring productive vocabulary (Mitchell et al., 2022). The importance of cognate status when learning new words is also supported by a study that taught 10-year-olds as well as college-aged adults words in a language unknown to them, some of them were cognates (Valente et al., 2018). They found better auditory recognition in both groups for cognates compared to non-cognates. Together, these results indicate that bilingual children might use cognate words as a “strategy” to boost their word learning, but any advantage of cognates in learning might no longer be present when the word is already learned and is just retrieved, at least in fairly balanced bilinguals. We can thus conclude that cognate status affects if and when a word is learned by young bilinguals (Bosch & Ramon-Casas, 2014; Mitchell et al., 2022; Schelletter, 2002) but does not seem to have strong effects on word recognition in this population. Further research is necessary to disentangle when and under what conditions cognate facilitation effects emerge.

4.4.3 Mispronunciation Sensitivity not Affected by Cognate Status

Our third prediction was that bilingual toddlers would be less sensitive to mispronunciations in cognates compared to non-cognates. Based on previous studies on Catalan-Spanish bilinguals (Ramon-Casas et al., 2009; Ramon-Casas & Bosch, 2010), we expected that, compared to non-cognates, cognates would be phonologically underspecified. This might occur because it is easier for children to encode all sounds in non-cognates, where each form is very distinct, compared to cognates where there is some overlap as well as subtle language-specific variation. We instead found that bilingual toddlers were similarly sensitive to mispronunciations in both cognates and non-cognates. Studies that have found differences between monolinguals and bilinguals in their phonological encoding have typically tested younger learners (e.g., Byers-Heinlein et al., 2013). Thus, it is possible that by age 24-36 months, bilingual children have already acquired sophisticated phonological encoding for many of their learned words.

However, our results seem to contrast with the findings of two studies that tested cognates and non-cognates respectively in Spanish-Catalan bilingual toddlers. The authors compared two types of cognates: those that contained a vowel mispronunciation that was meaningful only in Catalan (specifically the /e/-/ɛ/ contrast) and those that contained a vowel mispronunciation that was meaningful in both Spanish and Catalan (for example the /e/-/i/ and /e/-/a/ contrasts). Children noticed the Catalan-only mispronunciations when inside non-cognate words but not inside cognate words (Ramon-Casas et al., 2009; Ramon-Casas & Bosch, 2010). However, for other contrasts, they were able to notice the mispronunciations inside cognates (Ramon-Casas et al., 2009). Thus, our research finding fits an explanation where in general, cognates and non-cognates are represented in similar phonological detail, but there may be some exceptions for words that contain certain contrasts, such as those used only in one language. Future research should further investigate the cases where cognates and non-cognates differ, to find out whether there is a representational difference between them. This can help us understand how the properties of the words learned affect their representation.

4.4.4 Differences Between Bilinguals and Monolinguals in Recognizing Words

Going beyond word recognition, whether or not bilinguals and monolinguals differ in their word learning is a much studied question (De Houwer et al., 2014; Gonzalez-Barrero et al., 2020; Pearson et al., 1993). Our study did not set out to compare the vocabulary sizes of the two groups, however our study design does allow to compare bilingual and monolingual toddlers on their word recognition. We found no main effect of language group on toddlers' ability to recognize words. This comparison was not part of our original research plan and pre-registration. Yet, based on prior research on bilingual-monolingual differences it seems reasonable to assume that toddlers that are exposed only to the language of test, as monolinguals are, would perform better than bilinguals tested only in one of their languages. Furthermore, for about half of the bilingual children tested in our study the language of test (English) was the child's non-dominant language. Thus, we would expect worse recognition in bilinguals compared to monolinguals, but we found no difference. It might be that any differences that exist between bilingual and monolingual toddlers in their word recognition abilities are small, and that our current study is not adequately powered to detect these effects. Indeed, a large archival study investigating the impact of language exposure on word recognition found that language exposure is an important predictor in whether children recognize a word (Sander Montant et al., 2022). Thus, despite the current study finding no difference between bilinguals and monolinguals' word recognition, it is likely that there are subtle but important differences between bilinguals' and monolinguals' word recognition that our study was not designed to capture.

4.4.5 Implications and Future Directions

The present study opens several routes for future investigation. First, we found that for 24- to 31-month-old toddlers, the main variable affecting their word recognition was whether the word was correctly pronounced or not. Cognate status and language background did not show evidence of affecting word recognition. We specifically tested simultaneous bilinguals who we expected to have acquired the cognates and non-cognates in this study. However, it is possible that if we had tested a younger age group, these other variables may have played a larger role. Specifically, it would be interesting to test the role of cognate status in development, comparing simultaneous and sequential bilingual toddlers and adults.

Second, we found no evidence that the cross-language similarity found in cognates affects word recognition, while other studies with children and adults have found a cognate effect in different tasks. This suggests that cognate status can affect tasks like word learning, but once words are learned, cognate status does not affect retrieval of those words. A future study could test whether cognates are easier to learn in a lab study, where toddlers are taught novel cognates and non-cognates and their performance on a recognition task is measured. In this case, we would expect an effect of cognate status on word recognition in the test phase, with cognates being learned more easily than non-cognates.

Finally, we found that cognates are not phonologically underspecified in our sample of bilingual toddlers. We chose to test both vowel and consonant mispronunciations, as a first attempt to better understand how cognate status and phonological representations interact. However, cognates may vary in different ways on vowels and consonants, and thus manipulating vowels and consonants separately would allow us to test whether one or the other is underspecified in cognates, which we did not have the power to detect. Exploring these three areas for future research would bring important nuance to the question of how early word learning happens and the factors that affect bilingual word learning.

4.4.6 Conclusion

Overall, we found the same patterns of word recognition for bilingual and monolingual toddlers: both groups were able to recognize both correctly pronounced and mispronounced words, with a decrease in performance for mispronunciations characteristic of detailed phonological representations. Furthermore, children performed similarly for cognate and non-cognate words, showing that word recognition in bilingual infants was not affected by their other language, even when they heard their non-dominant language. Our findings further highlight that in word recognition, bilingual and monolingual children may be more similar than different, despite being exposed to very different learning environments.

5 Discussion

Across five studies described in three chapters, I investigated bilingual and monolingual children's ability to navigate their language input. This dissertation had two main research objectives. The first objective was to test infants' ability to discriminate languages in situations that are typical of bilingual infants' everyday life. The second objective was to test whether cross-language similarity affects toddlers' word representations. Both research objectives fit under the broader umbrella of exploring what young language learners are attentive to and what strategies they might be using to make sense of their language environment.

In Chapter 2 I asked whether young infants can detect language switches at short timescales, specifically single-word switches. I tested 8-12-month-old bilingual and monolingual infants on their ability to distinguish between single-language and switched-language utterances. Infants showed no evidence of distinguishing between the two types of trials when they heard language switches in a list of single words, and inconclusive evidence when hearing the switch in a sentence. No differences between language groups were found. Thus, we have no evidence that infants of this age can detect single-word language switches. In the chapter I discussed that these findings are important to add to the existing literature, as they highlight the need for a more nuanced narrative on infant language discrimination.

In Chapter 3 I asked whether infants make associations between a person and the language spoken by that person. Infants aged 5-18-months were familiarized with a man speaking English and a woman speaking French (or vice versa), and then tested on trials with the same pairing (man speaking English) and the switched pairing (man speaking French). Infants were tested either in an auditory-only condition (only hearing the speaker) or an audiovisual condition (both hearing and seeing the speaker). Bilingual and monolingual infants of all ages showed no difference in looking time to the two types of trials regardless of stimuli, and thus we have no evidence that infants are able to make person-language associations. This refutes the idea that the one-person-one-language rule for bilingual parenting is necessary for bilingual development, by showing that infants do not appear to form person-language associations rapidly and automatically at this age.

In Chapter 4 I asked whether bilingual exposure, and specifically cross-language similarity, affects toddlers' word representations. Two-year-olds were tested on their recognition of familiar cognates (e.g., “banana”, whose French translation is *banane*) and non-cognates (e.g., “apple”, whose French translation is *pomme*) in a looking-while-listening eye tracking study. Half of the words in the study were mispronounced, to probe how phonologically detailed toddlers' word representations were. I found that both bilingual and monolingual toddlers looked less at the target when it was mispronounced, indicating that they had detailed phonological representation of familiar words. There were no differences between cognate and non-cognate words, indicating that for simultaneous bilinguals, cross-linguistic similarity may not affect word recognition. I then explored how testing different groups of bilinguals or children of different ages affect how cross-language similarity interacts with language processing.

In the following sections I discuss the main contributions of the studies in this thesis, focussing specifically on language discrimination as well as word recognition in bilingual and monolingual children. Building on this, I discuss the broader implications of this work, including how and when bilingual-monolingual comparisons are informative about language learning in general, as well as the role of null results in advancing developmental science. I end the discussion by exploring limitations of the present research and possible future directions to expand our knowledge of language development.

5.1 Main Contributions

This thesis makes at least two main contributions. The first one, drawn from Chapters 2 and 3, concerns children’s ability to discriminate languages in situations that are typical in bilinguals’ everyday life. The studies in Chapters 2 and 3 indicate that the question of when children can discriminate languages is much more complex than the existing literature on language discrimination suggests. The second overarching contribution is derived from Chapter 4: bilingual and monolingual toddlers’ word representations are phonologically detailed, and also, word recognition and phonological encoding are not affected by how similar-sounding the word’s translation is. This suggests that the impact of cross-language similarity on word recognition for bilinguals is task dependent.

5.1.1 Language Discrimination

Previous research on infants’ ability to discriminate languages has focussed on testing language discrimination in long passages of speech that alternate. This research has demonstrated that some languages are discriminated earlier than others (Gasparini et al., 2021). For example, English and Japanese are discriminated at birth (Nazzi et al., 1998), but English and Dutch are only discriminated at 5 months of age (Nazzi et al., 2000). Furthermore, an important factor that explains why some languages are discriminated earlier than others is the duration cues of vowels and consonants present in long utterances of speech, which is often described as language rhythm (for review, see Gasparini et al., 2021). However, bilingual infants' language environments are much more nuanced than these language discrimination study paradigms. Long single-language passages do not comprise the entirety of bilinguals' experience. It is also typical for bilinguals to hear an utterance in one language with some words borrowed from their other language (Bail et al., 2015; Kremin et al., in press). This type of speech puts much higher demands on their ability for language discrimination as single words carry fewer rhythmic cues that help with language discrimination. And they hear multiple caregivers, some or all of them bilingual who each use their languages in an idiosyncratic manner. Thus, to the extent that infants can learn associations between a person and the language they speak, they can use this information to help them keep track of which languages are being spoken.

In Chapters 2 and 3 I showed that language discrimination is much more nuanced when testing infants in situations that might better reflect the entirety of bilingual infants' day-to-day language environments. While research using long passages of stimuli suggested that infants might discriminate languages easily after 5 months (Bosch & Sebastian Galles, 2001; Molnar et al., 2014; Nazzi et al., 2000), this was not the case when testing single-word language discrimination, as I did in Chapter 2. Similarly, if language discrimination is easy for infants to achieve, we would have expected infants to learn to associate a person and the language they speak, similar to how they learn other characteristics about people (Kinzler et al., 2010). Yet, in Chapter 3, I found that infants do not spontaneously make these language-person associations. This suggests that the timeline of when bilingual infants discriminate their languages is much more complex than previously assumed. Additional research is needed to uncover what happens between 5 months of age, which is the oldest age at which there have been language discrimination studies showing successful discrimination using long passages of speech, and 20 months, when infants are able to discriminate between single words from different languages in a comprehension task (Byers-Heinlein et al., 2017; Potter et al., 2018, 2019). Finding out when infants are first able to discriminate languages using single words or short phrases can illuminate how bilingual infants cope with their language input.

The results in this dissertation can be interpreted in the context of theories of language categorization (Byers-Heinlein, 2014). The language categorization framework differentiates "perceptual categorization", which is based on perceptual differences in the input, and "conceptual categorization", which builds on perceptual categorization but includes higher-level knowledge about the different language categories. Under this proposal, infants initially differentiate their languages through perceptual categorization, where low-level auditory information such as rhythm helps infants discriminate their languages. Over time, infants learn to treat these perceptual categories as conceptually different (e.g., as belonging to different languages), which is conceptual categorization. Thus, it may be that infants in their first year of life rely on perceptual categorization to discriminate languages, which is consistent with the literature that they can discriminate long passages of speech. In order to discriminate between single words from different languages, or in order to associate a person and the language they speak, infants might need to be able to achieve conceptual categorization, which may only develop in their second year of life. More research will be needed to fully test this idea. Overall, the findings from Chapters 2 and 3 highlight that infants' ability to discriminate languages is more nuanced and complex than what the existing research suggests.

5.1.2 Word Recognition

A major milestone in children's early years is when they start to understand words, for example, when a caregiver asks them if they would like a banana. This expands children's ability to interact with those around them and express their needs and wants. A particularity of bilingual children's language environment is that they typically learn two words for a concept, such as "apple" and "*pomme*" (fr. apple). These word pairs are called translation equivalents. Translation equivalents that sound similar across the two languages, such as "banana" and "*banane"* (fr. banana), are also called cognates. The overlap in sounds of cognates may make it easier for bilingual children to learn both forms, and this has indeed been documented in vocabulary composition studies (Mitchell et al., 2022; Schelletter, 2002). Furthermore, in some language processing tasks bilingual infants performed better on cognate compared to non-cognate trials (Pérez et al., 2010; Von Holzen et al., 2018). In Chapter 4 which tested infants on a word comprehension task, I found no difference between cognate and non-cognate trials, suggesting that some aspects of language processing might be unaffected by cross-language similarity.

An interesting moderator of whether cross-language similarity affects language processing may be the age when the tested bilingual populations acquired their second language. In our study, all children were exposed to both languages in the first year of life (most learned both from birth), whereas in von Holzen's study (2018) showing cognate effects in word recognition, the authors tested sequential bilinguals who started learning their second language at 20 months on average. The Revised Hierarchical Model (Kroll et al., 2010) can explain how cognate effects can play out differently for bilinguals who learned their second language at different ages. According to this model, for sequential bilinguals, who learned their second language after their first language, word recognition in their second language would go through the first language. Word recognition in the first language is faster than in the second language for these bilinguals, and for cognate words specifically word recognition can be sped up because of the phonological overlap between both forms. For simultaneous bilinguals on the other hand, the representations in both languages are equally fast, and no cognate effects are observed, potentially also because word recognition is very fast and automatic in both languages. A similar difference was reported in a study that tested "balanced" bilinguals, who had approximately equal exposure to both languages, and "imbalanced" bilinguals. In a picture naming task, cognate effects only showed up in the latter group but not in the former (Pérez et al., 2010). Overall, the results of Chapter 4 combined with the existing literature suggest that cognate effects may be moderated by the type and timing of bilingual exposure infants have received.

A second aspect of word recognition I tested was whether bilingual and monolingual toddlers were sensitive to mispronunciations, which would indicate that they have detailed phonological word representations. While most previous studies found that bilingual and monolingual children are sensitive to mispronunciations (Singh et al., 2020; Tamási et al., 2016; Wewalaarachchi et al., 2017), some reported that bilinguals did not notice subtle mispronunciations that were only contrastive in one of their languages (Ramon-Casas et al., 2009). In Chapter 4, I found that bilingual and monolingual toddlers were equally sensitive to mispronunciations in familiar words, consistent with most of the literature to date. This indicates further that bilingual toddlers do not have underspecified phonological word representations compared to monolinguals. This is also important in light of some of the studies showing that bilingual children sometimes have difficulties using subtle contrasts to guide their word learning (Fennell et al., 2007; Fennell & Byers-Heinlein, 2014; Havy et al., 2016). Overall, the findings in

Chapter 4 are in line with the existing research and also fit with the PRIMIR framework for language acquisition (Curtin et al., 2011). The PRIMIR framework posits that while in some cases, bilingual and monolingual performance on language tasks may diverge, in other cases bilinguals and monolinguals can perform similarly, even if different mechanisms may be used. The results from Chapter 4 show that bilingual exposure does not necessarily lead to an underspecified phonological representation in a word recognition task.

5.2 Broader Implications

In the following, I will discuss two broader implications of the present work for the field of language development and developmental science more generally. First, in Chapters 2 and 3 of this dissertation I report null results, and thus I discuss the role of null results in developmental research. Second, a throughline of the findings in this work was the absence of differences between bilingual and monolingual children's performance; I discuss how this informs how the field thinks about the value of comparisons between language groups.

5.2.1 Importance of Null Results

Two out of three empirical chapters in this dissertation reported null results for the main research questions: in Chapters 2 and 3, I did not find any evidence that infants could succeed in the task at hand. I also did not observe a difference between bilingual and monolingual groups, a point that I will discuss in more detail in the next section. The present set of manuscripts underlines why it is important to publish null results. Without the manuscripts contained in Chapters 2 and 3, the main takeaway regarding the published literature on newborn language discrimination would be that infants are able to discriminate between languages by 5 months. Yet, the null results I report contradict this categorical conclusion.

In many fields, null results are less likely to be published than positive results, because journals favour studies with statistically significant results and because researchers may be hesitant to invest time in publishing them (C. J. Ferguson & Heene, 2012). However, this can lead to systematic bias in the published literature, also called the "file drawer effect" (Rosenthal, 1979). In developmental research, where participant samples are hard to recruit and take a long time to collect, not publishing null results would lead to a waste of precious resources. Indeed, in a meta-analysis of meta-analyses in language acquisition, about 11% of effect sizes included in the meta-analyses originated from unpublished studies, and the inclusion of these unpublished effect sizes did not change the results of the meta-analyses (Tsuji et al., 2020). This suggests that the amount of publication bias in developmental research is low. This is potentially due to a tendency of infant researchers to publish all samples tested. Moreover, some null results in developmental research are expected, as in many cases children initially cannot succeed in a task, then at some later age they succeed. This allows researchers to determine the trajectory of skill building in children. Thus, it is important that thoughtfully designed studies, regardless of their results, are published, so that other researchers can adapt their paradigms and explore moderating variables.

I believe that the studies in Chapters 2 and 3 report true null results. In addition to null hypothesis significance testing against chance, I also used equivalence tests to determine how conclusive the null results were (Lakens et al., 2018) and found, in most cases, that the null hypothesis was supported. However, in infant research it is always possible that null results are due to a methodological reason, rather than the absence of a skill at a particular age. It is possible that, had I used a different task, infants may have succeeded in the studies in Chapters 2 and 3. However, all studies used carefully designed paradigms that had been used in previous studies for both Chapter 2 (Willits et al., 2013) as well as Chapter 3 (Weikum et al., 2007). I will elaborate on this point in the section on Limitations and Future Directions.

5.2.2 Bilingual-Monolingual Comparisons

The question of whether and which differences exist between bilingual and monolingual children has received a lot of interest in language acquisition research (Bialystok, 2021; De Houwer et al., 2014; Peal & Lambert, 1962; Poulin-Dubois et al., 2011). This question originally stemmed from a concern that bilingualism might be detrimental to children's language development (Peal & Lambert, 1962). However, in the past decades the question is used more as a vehicle to test how their lifelong language exposure affects bilingual and monolingual children differently (for a review, see Rocha-Hidalgo & Barr, in press). Each of the three empirical chapters in this thesis tested a monolingual control group, in addition to bilinguals. In Chapters 2 and 3, testing monolingual children allowed me to investigate whether lifelong bilingual exposure is necessary to succeed in the study task, specifically the task of detecting language switches (Chapter 2), and the task of associating a person with the language they speak (Chapter 3). In Chapter 4, I tested monolinguals as a control group to ensure that differences between cognates and non-cognates only affect bilinguals (who are the only group to comprehend both words) and are not an artefact of the words tested. In all three chapters, I found no differences between bilingual and monolingual groups.

The bilingual-monolingual comparison in this dissertation allowed me to elucidate whether and how lifelong language exposure affects language processing in children. However, bilingual-monolingual comparisons have sometimes been approached through a deficit lens where bilinguals are cast as “delayed”, or using a reductionistic lens where the description of bilingual experience is lacking the nuance it requires (López et al., 2021). Consequently, in the language acquisition literature, bilingual-monolingual comparisons are somewhat controversial. A famous quotation states “The bilingual is not two monolinguals in one person” (Grosjean, 1989), highlighting that viewing bilinguals through a monolingual lens will always give an incomplete picture of the bilingual experience. Furthermore, comparisons between bilinguals and monolinguals often assume that monolinguals are the ‘default’, which is not supported when looking at the numbers of bilinguals versus monolinguals worldwide. Using monolingual populations as the ‘default’ is furthermore problematic as it is a western-centric perspective on language norms. My studies did not report any bilingual-monolingual differences, highlighting the developmental similarities between these groups. Moreover, in the future, it would be interesting to explore continuous approaches to studying bilingual and monolingual experiences (Kremin & Byers-Heinlein, 2021).

However, it is important to keep in mind that we cannot conclude from observing the same behaviour that the same underlying processes lead to that behaviour in both populations. An example in this thesis would be bilingual and monolingual children’s similar sensitivity to mispronunciations in Chapter 4. It may be that the correspondence between sounds and concepts is represented differently in bilingual and monolingual children, and in this case, we could still observe that both groups are sensitive to mispronunciations to a similar degree. Nevertheless, testing groups of bilinguals and monolinguals is important to help disentangle which behaviours are only observed following a lifelong exposure to multiple languages, and which ones children of all language backgrounds exhibit spontaneously. This can further inform us about the strategies bilingual children use to adapt to their language environment.

5.3 Limitations and Future Directions

The three empirical chapters in this thesis contributed important insights about bilingual and monolingual development, but the design choices I made for each study necessarily limit some aspects of the conclusions while providing directions for future research. A limitation of the present work is that the age groups I chose to test, as well as the sample sizes, were limited by available resources.

In terms of sample sizes, the groups in each study had an average sample size of 19-25 participants. In each study I used a sensitivity analysis to determine the smallest effect size that I had 80% power to detect. The median sample size hovers around 18 participants per cell in developmental psychology (Bergmann et al., 2018; Oakes, 2017), thus my sample sizes were either comparable or larger than those of other studies in the field. Language acquisition research often has underpowered study designs (Bergmann et al., 2018), likely in part because testing developmental populations is slow and uses a lot of resources (Frank et al., 2017). Recruiting participants is a challenge for developmental studies (Schott et al., 2019). This challenge is exacerbated for recruiting bilingual participants, where time-consuming structured interviews are necessary to determine participant eligibility (Byers-Heinlein et al., 2020; DeAnda et al., 2016) to ensure validity. In an ideal world I would have tested more participants, but in actuality I tested a number of participants that was reasonable in order to draw appropriately powered statistical conclusions while working under the constraints of limited time and resources.

In each study I carefully chose specific ages to test given the previous literature related to my research questions. However, it is possible that children a little bit older or a little bit younger would have shown an interesting developmental trajectory that my samples were not able to capture. Just as for increasing sample sizes, recruiting infants in additional age groups would have multiplied the resources required to recruit, test participants, and clean data. These were already extensive for the ages of participants tested in the current studies. In Chapters 2 (single-word language discrimination) and 4 (word recognition) I tested one age group that spanned several months, and I did not have a sufficient sample size to investigate age-related individual differences within each sample. Regarding the age groups studied in Chapter 2, future studies could test older children to elucidate the trajectory from 12 months to 20 months, when children start to detect single-word language switching in a comprehension task. In Chapter 3, I tested 5-, 12-, and 18-month-olds and did not find that children made the association under study at any age, however it is possible that older children would succeed in this task. Regarding this study, future research could investigate 3–4-year-olds' ability to make language-person associations. At this age, children's language choice in response to adults speaking different languages could also be tested (Genesee et al., 1996). This would allow us to draw parallels between the perceptual underpinnings of language production. For Chapter 4, future studies could test younger children to assess whether cognate and mispronunciation effects are stronger in this group. Younger children have weaker language skills and thus may rely more on strategies such as cross-language similarity, or they may have underspecified representation of sounds. In summary, widening the age groups being studied and increasing sample sizes would help provide more detailed answers on bilingual language acquisition.

5.4 Conclusion

How children learn language by making sense of their language environment is a complex question. In this dissertation, I showed that infants do not show evidence of discriminating languages in some tasks, calling into question how salient the differences between languages are at this age. Bilingual and monolingual infants did not detect single-word language switches (Chapter 2), nor did they learn person-language associations (Chapter 3). During toddlerhood when word learning is in full swing, I found that bilingual and monolingual toddlers encode phonological detail for familiar words, regardless of the cross-language similarity (Chapter 4). This work highlights the need for careful examination of the cues that bilingual children use to acquire each of their languages, as well as how the outcome of learning is different or similar between bilingual and monolingual children.

References

Albareda-Castellot, B., Pons, F., & Sebastian Galles, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm: Acquisition of phonetic categories in bilingual infants. *Developmental Science*, *14*(2), 395–401. https://doi.org/10.1111/j.1467-7687.2010.00989.x

Andras, F., Rivera, M., Bajo, T., Dussias, P. E., & Paolieri, D. (2022). Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study. *International Journal of Bilingualism*, 13670069211033360. https://doi.org/10.1177/13670069211033359

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Assche, E. V., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00174

Aust, F. (2019). *citr: "RStudio" add-in to insert markdown citations* (0.3.2). https://github.com/crsh/citr

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown* (0.1.0.9997). https://github.com/crsh/papaja

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory Redundancy Guides the Development of Selective Attention, Perception, and Cognition in Infancy. *Current Directions in Psychological Science*, *13*(3), 99–102. https://doi.org/10.1111/j.0963-7214.2004.00283.x

Bahrick, L. E., & Pickens, J. N. (1988). Classification of bimodal English and Spanish language passages by infants. *Infant Behavior and Development*, *11*(3), 277–296. https://doi.org/10.1016/0163-6383(88)90014-8

Bail, A., Morini, G., & Newman, R. S. (2015). Look at the gato! Code-switching in speech to toddlers. *Journal of Child Language*, *42*(05), 1073–1101. https://doi.org/10.1017/S0305000914000695

Barron-Hauwaert, S. (2004). The One-Parent-One-Language Approach. What is it? In *Language Strategies for Bilingual Families: The one-parent-one-language Approach* (pp. 1–22). Multilingual Matters. https://doi.org/doi:10.21832/9781853597169-002

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369–380. https://doi.org/10.1080/15475441.2014.979387

Bergelson, E., & Swingley, D. (2018). Young Infants' Word Comprehension Given An Unfamiliar Talker or Altered Pronunciations. *Child Development*, *89*(5), 1567–1576. https://doi.org/10.1111/cdev.12888

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Bialystok, E. (2021). Cognitive effects of bilingualism: An evolving perspective. In W. S. Francis (Ed.), *Bilingualism Across the Lifespan* (pp. 9–28). Routledge.

Birulés, J., Pons, F., & Bosch, L. (2018). *Detection of a language switch from a talking face: Evidence from monolingual and bilingual infants.* Poster presented at the International Congress of Infant Studies (ICIS), Philadelphia, USA.

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (6.0.22). http://www.praat.org/

Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, *38*(4), 317–322. https://doi.org/10.1177/0165025414532559

Bosch, L., & Sebastian Galles, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49. https://doi.org/10.1207/S15327078IN0201_3

Bosma, E., & Nota, N. (2020). Cognate facilitation in Frisian–Dutch bilingual children's sentence reading: An eye-tracking study. *Journal of Experimental Child Psychology*, *189*, 104699. https://doi.org/10.1016/j.jecp.2019.104699

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–B44. https://doi.org/10.1016/S0010-0277(01)00122-6

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*(03). https://doi.org/10.1017/S0142716407070257

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-old infants. *Child Development*, *84*(2), 422–428. https://doi.org/10.1111/j.1467-8624.2012.01860.x

Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition*, *16*(01), 32–48. https://doi.org/10.1017/S1366728912000120

Byers-Heinlein, K. (2014). Languages as categories: Reframing the "one language or two" question in early bilingual development. *Language Learning*, *64*(s2), 184–201. https://doi.org/10.1111/lang.12055

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*(3), 343–348. https://doi.org/10.1177/0956797609360758

Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology*, *56*(2), 274–291. https://doi.org/10.1002/dev.21167

Byers-Heinlein, K., Fennell, C. T., & Werker, J. F. (2013). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition*, *16*(01), 198–205. https://doi.org/10.1017/S1366728912000417

Byers-Heinlein, K., Jardak, A., Fourakis, E., & Lew-Williams, C. (2022). Effects of language mixing on bilingual children's word learning. *Bilingualism: Language and Cognition*, *25*(1). https://doi.org/10.1017/S1366728921000699

Byers-Heinlein, K., Morin-Lessard, E., & Lew-Williams, C. (2017). Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(34), 9032–9037. https://doi.org/10.1073/pnas.1703220114

Byers-Heinlein, K., Schott, E., Gonzalez-Barrero, A. M., Brouillard, M., Dubé, D., Jardak, A., Laoun-Rubenstein, A., Mastroberardino, M., Morin-Lessard, E., Iliaei, S. P., Salama-Siroishka, N., & Tamayo, M. P. (2020). MAPLE: A multilingual approach to parent language estimates. *Bilingualism: Language and Cognition*, *23*(5), 951–957. https://doi.org/10.1017/S1366728919000282

Carbajal, M. J., Peperkamp, S., & Tsuji, S. (2021). A meta-analysis of infants' word-form recognition. *Infancy*, *26*(3), 369–387. https://doi.org/10.1111/infa.12391

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521–536. https://doi.org/10.1037/dev0000083

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*(4), 492–504. https://doi.org/10.1016/j.wocn.2010.12.002

De Groot, A. M. B., & Keijzer, R. (2000). What Is Hard to Learn Is Easy to Forget: The Roles of Word Concreteness, Cognate Status, and Word Frequency in Foreign-Language Vocabulary Learning and Forgetting. *Language Learning*, *50*(1), 1–56. https://doi.org/10.1111/0023-8333.00110

de Haan, M., Johnson, M. H., Maurer, D., & Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. *Cognitive Development*, *16*(2), 659–678. https://doi.org/10.1016/S0885-2014(01)00051-X

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, *28*(03). https://doi.org/10.1017/S0142716407070221

De Houwer, A., Bornstein, M. H., & Putnick, D. (2014). A bilingual-monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, *35*, 1189–1211. https://doi.org/doi:10.1017/S0142716412000744

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). The Language Exposure Assessment Tool: Quantifying Language Exposure in Infants and Children. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1346–1356. https://doi.org/10.1044/2016_JSLHR-L-15-0234

Dehaene-Lambertz, G., & Houston, D. (1998). Faster Orientation Latencies Toward Native Language in Two-Month-Old Infants. *Language and Speech*, *41*(1), 21–43. https://doi.org/10.1177/002383099804100102

Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language*, *72*, 1–15. https://doi.org/10.1016/j.jml.2013.12.001

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03), 175–197. https://doi.org/10.1017/S1366728902003012

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fawcett, C. (2021). Sharing a Common Language Affects Infants' Pupillary Contagion. *Journal of Cognition and Development*, *23*(2), 173–187. https://doi.org/10.1080/15248372.2021.2013225

Fecher, N., & Johnson, E. K. (2018). The native-language benefit for talker identification is robust in 7.5-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1911–1920. https://doi.org/10.1037/xlm0000555

Fecher, N., & Johnson, E. K. (2019). Bilingual infants excel at foreign-language talker recognition. *Developmental Science*, *22*(4), e12778. https://doi.org/10.1111/desc.12778

Fecher, N., & Johnson, E. K. (2022). Revisiting the talker recognition advantage in bilingual infants. *Journal of Experimental Child Psychology*, *214*, 105276. https://doi.org/10.1016/j.jecp.2021.105276

Fennell, C. T., & Byers-Heinlein, K. (2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *International Journal of Behavioral Development*, *38*(4), 309–316. https://doi.org/10.1177/0165025414530631

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using Speech Sounds to Guide Word Learning: The Case of Bilingual Infants. *Child Development*, *78*(5), 1510–1525. https://doi.org/10.1111/j.1467-8624.2007.01080.x

Fenson, L., Marchman, V. A., & Thal, D. (2007). *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual*. Paul H. Brookes Publishing Company.

Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports*, *6*, 25434. https://doi.org/10.1038/srep25434

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Floccia, C., Nazzi, T., & Bertoncini, J. (2000). Unfamiliar voice discrimination for short stimuli in newborns. *Developmental Science*, *3*(3), 333–343. https://doi.org/10.1111/1467-7687.00128

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one- to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, *41*(05), 1085–1114. https://doi.org/10.1017/S0305000913000287

Forbes, S. H. (2020). PupillometryR: An R package for preparing and analysing pupillometry data. *Journal of Open Source Software*, *5*(50), 2285. https://doi.org/10.21105/joss.02285

Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., Hamlin, J. K., Hannon, E. E., Kline, M., Levelt, C., Lew-Williams, C., Nazzi, T., Panneton, R., Rabagliati, H., Soderstrom, M., Sullivan, J., Waxman, S., & Yurovsky, D. (2017). A Collaborative Approach to Infant Research: Promoting Reproducibility, Best Practices, and Theory-Building. *Infancy*, *22*(4), 421–435. https://doi.org/10.1111/infa.12182

Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language*, *89*, 110–137. https://doi.org/10.1016/j.jml.2015.10.001

Fritzsche, T., & Höhle, B. (2015). Phonological and lexical mismatch detection in 30-month-olds and adults measured by pupillometry. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)* (p. 5). University of Glasgow.

Gasparini, L., Langus, A., Tsuji, S., & Boll-Avetisyan, N. (2021). Quantifying the role of rhythm in infants’ language discrimination abilities: A meta-analysis. *Cognition*, *213*, 104757. https://doi.org/10.1016/j.cognition.2021.104757

Genesee, F., Boivin, I., & Nicoladis, E. (1996). Talking with strangers: A study of bilingual children’s communicative competence. *Applied Psycholinguistics*, *17*(4), 427–442. https://doi.org/10.1017/S0142716400008183

Gonzalez-Barrero, A. M., Ben, R. D., Killam, H., & Byers-Heinlein, K. (2021). *Word learning in the lab is hard! Data from 14-month-old monolingual and bilingual infants*. PsyArXiv. https://doi.org/10.31234/osf.io/5dvx8

Gonzalez-Barrero, A. M., Schott, E., & Byers-Heinlein, K. (2020). *Bilingual adjusted vocabulary: A developmentally-informed bilingual vocabulary measure*. PsyArXiv. https://doi.org/10.31234/osf.io/x7s4u

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the Rhythm Class Hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 515–546). De Gruyter. https://doi.org/10.1515/9783110197105.2.515

Grammont, M. (1902). Observations sur le langage des enfants. In D. Barbelent, G. Dottin, R. Gauthiot, M. Grammont, Laronde, M. Niedermann, & J. Vendreyes (Eds.), *Mélanges linguistiques. Offerts à M. Antoine Meillet par ses élèves* (pp. 61–82). Klincksieck.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Havy, M., Bouchon, C., & Nazzi, T. (2016). Phonetic processing when learning words: The case of bilingual infants. *International Journal of Behavioral Development*, *40*(1), 41–52. https://doi.org/10.1177/0165025415570646

Hillairet de Boisferon, A., Dupierrix, E., Quinn, P. C., Lœvenbruck, H., Lewkowicz, D. J., Lee, K., & Pascalis, O. (2015). Perception of Multisensory Gender Coherence in 6- and 9-Month-Old Infants. *Infancy*, *20*(6), 661–674. https://doi.org/10.1111/infa.12088

Hirschfeld, L. A., & Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, *12*(2), 213–238. https://doi.org/10.1016/S0885-2014(97)90014-9

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, *14*(6), 1445–1458. https://doi.org/10.1111/j.1467-7687.2011.01089.x

Howard, L. H., Carrazza, C., & Woodward, A. L. (2014). Neighborhood Linguistic Diversity Predicts Infants' Social Learning. *Cognition*, *133*(2), 474–479. https://doi.org/10.1016/j.cognition.2014.08.002

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience: Infant voice discernment. *Developmental Science*, *14*(5), 1002–1011. https://doi.org/10.1111/j.1467-7687.2011.01052.x

Jørgensen, R. N., Dale, P. S., Bleses, D., & Fenson, L. (2010). CLEX: A cross-linguistic lexical norms database*. *Journal of Child Language*, *37*(2), 419–428. https://doi.org/10.1017/S0305000909009544

Kandhadai, P., Danielson, D. K., & Werker, J. F. (2014). Culture as a binder for bilingual acquisition. *Trends in Neuroscience and Education*, *3*(1), 24–27. https://doi.org/10.1016/j.tine.2014.02.001

Kartushina, N., & Mayor, J. (2019). Word knowledge in six- to nine-month-old Norwegian infants? Not without additional frequency cues. *Royal Society Open Science*, *6*(9), 180711. https://doi.org/10.1098/rsos.180711

Kinzler, K. D. (2021). Language as a Social Cue. *Annual Review of Psychology*, *72*(1), 241–264. https://doi.org/10.1146/annurev-psych-010418-103034

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577–12580. https://doi.org/10.1073/pnas.0705345104

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*(4), 581–592. https://doi.org/10.1002/ejsp.739

Kremin, L. V., & Byers-Heinlein, K. (2021). Why not both? Rethinking categorical and continuous approaches to bilingualism. *International Journal of Bilingualism*, *25*(6), 1560–1575. https://doi.org/10.1177/13670069211031986

Kremin, L. V., Orena, A. J., Polka, L., & Byers-Heinlein, K. (in press). Code-switching in parents' everyday speech to bilingual infants. *Journal of Child Language*. https://doi.org/10.1017/S0305000921000118

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(03), 373–381. https://doi.org/10.1017/S136672891000009X

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. https://doi.org/10.1111/j.1467-7687.2006.00468.x/full

Kuipers, J.-R., & Thierry, G. (2012). Event-related potential correlates of language change detection in bilingual toddlers. *Developmental Cognitive Neuroscience*, *2*(1), 97–102. https://doi.org/10.1016/j.dcn.2011.08.002

Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lalonde, K., & Werner, L. A. (2019). Infants and Adults Use Visual Cues to Improve Detection and Discrimination of Speech in Noise. *Journal of Speech, Language, and Hearing Research : JSLHR*, *62*(10), 3860–3875. https://doi.org/10.1044/2019_JSLHR-H-19-0106

Lewkowicz, D., & Tift, A. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 1431–1436. https://doi.org/10.1073/pnas.1114783109

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, *41*(Suppl 3), 622–634. https://doi.org/10.1111/cogs.12403

López, B. G., Luque, A., & Piña-Watson, B. (2021). Context, intersectionality, and resilience: Moving toward a more holistic study of bilingualism in cognitive science. *Cultural Diversity and Ethnic Minority Psychology*. https://doi.org/10.1037/cdp0000472

Martin, C. D., Molnar, M., & Carreiras, M. (2016). The proactive bilingual brain: Using interlocutor identity to generate predictions for language processing. *Scientific Reports*, *6*. https://doi.org/10.1038/srep26171

Mattock, K., Polka, L., Rvachew, S., & Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: Comparing monolingual and bilingual infants. *Developmental Science*, *13*(1), 229–243. https://doi.org/10.1111/j.1467-7687.2009.00891.x

May, L., Baron, A. S., & Werker, J. F. (2019). Who can speak that language? Eleven-month-old infants have language-dependent expectations regarding speaker ethnicity. *Developmental Psychobiology*, *61*(6), 859–873. https://doi.org/10.1002/dev.21851

MetaLab Project. (2020). Retrieved from http://metalab.stanford.edu/

Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. *Infant Behavior and Development*, *6*(2), 313–330. https://doi.org/10.1016/S0163-6383(83)80040-X

Mitchell, L., Tsui, R. K. Y., & Byers-Heinlein, K. (2022). *Cognates are advantaged in early bilingual expressive vocabulary development*. PsyArXiv. https://doi.org/10.31234/osf.io/daktp

Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2014). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism: Language and Cognition*, *17*(3), 445–463. https://doi.org/10.1017/S1366728913000576

Molnar, M., & Davidson, D. (2017). *Do bilingual 8-month-olds care when their caregivers break the one parent, one language rule?* Poster presented at the Workshop on Infant Language Development, Bilbao, Spain.

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. https://doi.org/10.1111/infa.12041

Moore, C., & Bergelson, E. (2021). *Examining the roles of regularity and lexical class in 18—26-month-olds' representations of how words sound*. OSF Preprints. https://doi.org/10.31219/osf.io/kp7tv

Morini, G., & Newman, R. S. (2019). Dónde está la ball? Examining the effect of code switching on bilingual children's word recognition. *Journal of Child Language*, *46*(6), 1238–1248. https://doi.org/10.1017/S0305000919000400

Morin-Lessard, E., Poulin-Dubois, D., Segalowitz, N., & Byers-Heinlein, K. (2019). Selective attention to the mouth of talking faces in monolinguals and bilinguals aged 5 months to 5 years. *Developmental Psychology*, *55*(8), 1640–1655. https://doi.org/10.1037/dev0000750

Muylle, M., Assche, E. V., & Hartsuiker, R. J. (2022). Comparing the cognate effect in spoken and written second language word production. *Bilingualism: Language and Cognition*, *25*(1), 93–107. https://doi.org/10.1017/S1366728921000444

Nacar Garcia, L., Guerrero-Mosquera, C., Colomer, M., & Sebastian-Galles, N. (2018). Evoked and oscillatory EEG activity differentiates language discrimination in young monolingual and bilingual infants. *Scientific Reports*, *8*(1), 2770. https://doi.org/10.1038/s41598-018-20824-0

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. https://doi.org/10.1037//0096-1523.24.3.756

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. https://doi.org/10.1006/jmla.2000.2698

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, *69*(1), 511–534. https://doi.org/10.1146/annurev-psych-122216-011836

Nencheva, M. L., Piazza, E. A., & Lew-Williams, C. (2021). The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning. *Developmental Science*, *24*(1), e12997. https://doi.org/10.1111/desc.12997

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1016/0364-0213(90)90024-Q

Oakes, L. M. (2017). Sample Size, Statistical Power, and False Conclusions in Infant Looking-Time Research. *Infancy*, *22*(4), 436–469. https://doi.org/10.1111/infa.12186

Olson, R. (2017). *Wisconsin Infant Studies Program. Retrieved from https://bitbucket.org/rholson1/wisp/src/default/* [Custom Matlab program]. https://bitbucket.org/rholson1/wisp/src/default/

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the language environment analysis recording system in analyzing French-English bilingual speech. *Journal of Speech, Language, and Hearing Research: JSLHR*, *62*(7), 2491–2500. https://doi.org/10.1044/2019_JSLHR-L-18-0342

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2020). What do bilingual infants actually hear? Evaluating measures of language input to bilingual-learning 10-month-olds. *Developmental Science*, *23*(2), e12901. https://doi.org/10.1111/desc.12901

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, *76*(27), 1–23. https://doi.org/10.1037/h0093840

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*(1), 93–120. https://doi.org/10.1111/j.1467-1770.1993.tb00174.x/full

Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services*, *4*(1), 55–67.

Poltrock, S., & Nazzi, T. (2015). Consonant/vowel asymmetry in early word form recognition. *Journal of Experimental Child Psychology*, *131*, 135–148. https://doi.org/10.1016/j.jecp.2014.11.011

Potter, C. E., Fourakis, E., Morin-Lessard, E., Byers-Heinlein, K., & Lew-Williams, C. (2018). Bilingual infants process mixed sentences differently in their two languages. In C. Kalish, M. Rau, T. Rogers, & X. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 900–905). Cognitive Science Society.

Potter, C. E., Fourakis, E., Morin-Lessard, E., Byers-Heinlein, K., & Lew-Williams, C. (2019). Bilingual toddlers' comprehension of mixed sentences is asymmetrical across their two languages. *Developmental Science*, *22*(4), e12794. https://doi.org/10.1111/desc.12794

Potter, C. E., & Lew-Williams, C. (2019). Infants' selective use of reliable cues in multidimensional language input. *Developmental Psychology*, *55*(1), 1–8. https://doi.org/10.1037/dev0000610

Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology*, *108*(3), 567–579. https://doi.org/10.1016/j.jecp.2010.10.009

R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Ramon-Casas, M., & Bosch, L. (2010). Are Non-Cognate Words Phonologically Better Specified than Cognates in the Early Lexicon of Bilingual Children? In M. Ortega-Llebaria (Ed.), *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 31–36). Cascadilla Proceedings Project.

Ramon-Casas, M., Fennell, C. T., & Bosch, L. (2017). Minimal-pair word learning by bilingual toddlers: The Catalan /e/-/ɛ/ contrast revisited. *Bilingualism: Language and Cognition*, *20*(3), 649–656. https://doi.org/10.1017/S1366728916001115

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. https://doi.org/10.1016/j.cogpsych.2009.02.002

Ramus, F. (2002). Acoustic correlates of linguistic rhythm: Perspectives. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2002* (pp. 115–120). Université de Provence. http://cogprints.org/2273/

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265–292. https://doi.org/10.1016/S0010-0277(00)00101-3

Rocha-Hidalgo, J., & Barr, R. (in press). Defining Bilingualism in Infancy and Toddlerhood: A Scoping Review. In *International Journal of Bilingualism*. https://doi.org/10.31234/osf.io/bmdhx

Rogers, J., Webb, S., & Nakata, T. (2015). Do the cognacy characteristics of loanwords make them more easily learned than noncognates? *Language Teaching Research*, *19*(1), 9–27.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Sander Montant, A., Lopez Perez, M., & Byers-Heinlein, K. (2022). *The more they hear the more they learn? Using data from bilinguals to test models of early lexical development*. PsyArXiv. https://doi.org/10.31234/osf.io/zd3m8

Santolin, C., Garcia-Castro, G., Zettersten, M., Sebastian Galles, N., & Saffran, J. R. (2021). Experience with research paradigms relates to infants' direction of preference. *Infancy*, *26*(1), 39–46. https://doi.org/10.1111/infa.12372

Schelletter, C. (2002). The effect of form similarity on bilingual children's lexical development. *Bilingualism: Language and Cognition*, *5*(2), 93–107. https://doi.org/10.1017/S1366728902000214

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157–166. https://doi.org/10.1017/S1366728910000623

Schott, E., Kremin, L. V., & Byers-Heinlein, K. (2022). The youngest bilingual Canadians: Insights from the 2016 Census regarding children aged 0-9. In *Canadian Public Policy* (Vol. 48, Issue 2). https://doi.org/10.31234/osf.io/6q9jg

Schott, E., Mastroberardino, M., Fourakis, E., Lew-Williams, C., & Byers-Heinlein, K. (2021). Fine-tuning language discrimination: Bilingual and monolingual infants' detection of language switching. *Infancy*, *26*(6), 1037–1056. https://doi.org/10.1111/infa.12429

Schott, E., Rhemtulla, M., & Byers-Heinlein, K. (2019). Should I test more babies? Solutions for transparent data peeking. *Infant Behavior and Development*, *54*, 166–176. https://doi.org/10.1016/j.infbeh.2018.09.010

Sebastian Galles, 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. https://doi.org/10.1177/0956797612436817

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

Singh, L., Fu, C. S., Tay, Z. W., & Golinkoff, R. M. (2018). Novel word learning in bilingual and monolingual infants: Evidence for a bilingual advantage. *Child Development*, *89*(3), e183–e198.

Singh, L., Tan, A., & Quinn, P. C. (2021). Infants recognize words spoken through opaque masks but not through clear masks. *Developmental Science*, *24*(6), e13117. https://doi.org/10.1111/desc.13117

Singh, L., Tan, A. R. Y., Lee, K., & Quinn, P. C. (2020). Sensitivity to race in language comprehension in monolingual and bilingual infants. *Journal of Experimental Child Psychology*, *199*, 104933. https://doi.org/10.1016/j.jecp.2020.104933

Sirois, S. (2018). *Pupillometry Walkthrough*. https://oraprdnt.uqtr.uquebec.ca/pls/public/gscw031?owa_no_site=314&owa_no_fiche=3

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Squires, L. R., Ohlfest, S. J., Santoro, K. E., & Roberts, J. L. (2020). Factors Influencing Cognate Performance for Young Multilingual Children's Vocabulary: A Research Synthesis. *American Journal of Speech-Language Pathology*, *29*(4), 2170–2188. https://doi.org/10.1044/2020_AJSLP-19-00167

Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition*, *106*(1), 234–258. https://doi.org/10.1016/j.cognition.2007.01.011

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, *39*(4), 505–513. https://doi.org/10.1016/j.wocn.2010.08.006

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, *8*(5), 432–443. https://doi.org/10.1111/j.1467-7687.2005.00432.x

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*(2), 252–269. https://doi.org/10.1016/j.jml.2008.11.003

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., & Höhle, B. (2017). Pupillometry registers toddlers' sensitivity to degrees of mispronunciation. *Journal of Experimental Child Psychology*, *153*(Supplement C), 140–148. https://doi.org/10.1016/j.jecp.2016.07.014

Tamási, K., Mckean, C., Gafos, A., & Höhle, B. (2019). Children's gradient sensitivity to phonological mismatch: Considering the dynamics of looking behavior and pupil dilation. *Journal of Child Language*, *46*(1), 1–23. https://doi.org/10.1017/S0305000918000259

Tamási, K., Wewalaarachchi, T., Hoehle, B., & Singh, L. (2016). *Measuring sensitivity to phonological detail in monolingual and bilingual infants using pupillometry*.

Tobii® Technology. (2010). *Tobii T60 XL Eye Tracker*. https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf

Trudeau, N., Frank, I., & Poulin-Dubois, D. (1999). Une adaptation en français québécois du McArthur Communicative Development Inventory. *Canadian Journal of Speech-Language Pathology and Audiology*, *22*, 151–163.

Tsui, A. S. M., Byers-Heinlein, K., & Fennell, C. T. (2019). Associative word learning in infancy: A meta-analysis of the switch task. *Developmental Psychology*, *55*(5), 934–950. https://doi.org/10.1037/dev0000699

Tsuji, S., Cristia, A., Frank, M. C., & Bergmann, C. (2020). Addressing publication bias in meta-analysis: Empirical findings from community-augmented meta-analyses of infant language development. *Zeitschrift Für Psychologie*, *228*(1), 50. https://doi.org/10.1027/2151-2604/a000393

Valente, D., Ferré, P., Soares, A., Rato, A., & Comesaña, M. (2018). Does phonological overlap of cognate words modulate cognate acquisition and processing in developing and skilled readers? *Language Acquisition*, *25*(4), 438–453. https://doi.org/10.1080/10489223.2017.1395029

Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English–Welsh bilingual infants. *Applied Psycholinguistics*, *28*(3), 475–493. https://doi.org/10.1017/S0142716407070269

Von Holzen, K., & Bergmann, C. (2021). The development of infants' responses to mispronunciations: A meta-analysis. *Developmental Psychology*, *57*(1), 1–18. https://doi.org/10.1037/dev0001141

Von Holzen, K., Fennell, C. T., & Mani, N. (2018). The impact of cross-language phonological overlap on bilingual and monolingual toddlers' word recognition. *Bilingualism: Language and Cognition*, *22*(3), 476–499. https://doi.org/10.1017/S1366728918000597

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' Bimodal Perception of Gender. *Ecological Psychology*, *3*(2), 55–75. https://doi.org/10.1207/s15326969eco0302_1

Weatherhead, D., Kandhadai, P., Hall, D. G., & Werker, J. F. (2021). Putting Mutual Exclusivity in Context: Speaker Race Influences Monolingual and Bilingual Infants' Word-Learning Assumptions. *Child Development*, *92*(5), 1735–1751. https://doi.org/10.1111/cdev.13626

Weatherhead, D., & White, K. S. (2016). He Says Potato, She Says Potahto: Young Infants Track Talker-Specific Accents. *Language Learning and Development*, *12*(1), 92–103. https://doi.org/10.1080/15475441.2015.1024835

Weatherhead, D., & White, K. S. (2018). And then I saw her race: Race-based expectations affect infants' word processing. *Cognition*, *177*, 87–97. https://doi.org/10.1016/j.cognition.2018.04.004

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. https://doi.org/10.1126/science.1137686

Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, *39*(4), 703–728. https://doi.org/10.1017/S0142716418000152

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. https://doi.org/10.1016/S0163-6383(84)80022-3

Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers. *Journal of Experimental Child Psychology*, *159*, 16–33. https://doi.org/10.1016/j.jecp.2017.01.009

Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, *18*(6), 1053–1075. https://doi.org/10.1111/infa.12026

Appendices

Appendix A: Consent Form

CONSENT TO PARTICIPATE IN THE MONOLINGUAL AND BILINGUAL DEVELOPMENT PROJECT

I understand that I have been asked to participate in a program of research being conducted by Dr. Krista Byers-Heinlein of the Centre for Research in Human Development and the Psychology Department of Concordia University, 514-848-2424 x2208, k.byers@concordia.ca

A. PURPOSE

I have been informed that the purpose of the research is to understand how children develop their language and conceptual skills.

B. PROCEDURES

I understand that my child's participation in the study will take approximately 10 minutes, and that my participation will take approximately 60 to 90 minutes. My child will be seated comfortably in a study room, and I or a caregiver designated by me will accompany my child at all times. My child will see an audio-visual presentation including one or more of the following: language sounds, non-language sounds, colourful pictures, or a live interaction with a researcher. My child's reactions throughout the study will be recorded on video and/or via an eye tracker, and will be kept by the researcher for future reference. I may be asked to complete questionnaires regarding my child's background, experience, and knowledge.

I understand that any data will be stored in a secure location at Concordia University, and that I will only be identified by code number to protect my confidentiality. I understand that this consent form and my contact information will be stored for 5 years. My questionnaires, videos, and other anonymous data will be stored for at least 5 years, following the Tri-Agency Statement of Principles on Digital Data Management (Government of Canada). When my data are no longer needed, paper copies will be shredded, and digital information will be erased.

C. RISKS AND BENEFITS

I understand that there are no known risks to participation in this study. As a thank you for my participation, I will receive a small gift for my child and a certificate.

D. CONDITIONS OF PARTICIPATION

- I understand that I am free to withdraw my consent and discontinue my and my child's participation at anytime without negative consequences.
- I understand that my and my child's participation in this study is CONFIDENTIAL.
- I understand that the data from this study may be published.

I HAVE CAREFULLY STUDIED THE ABOVE AND UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND VOLUNTARILY AGREE TO PARTICIPATE WITH MY CHILD IN THIS STUDY.

CHILD'S NAME (please print) ____

PARENT'S NAME (please print) ____

SIGNATURE ____

DATE ____

I would be interested in participating in other studies conducted through the Centre for Research in Human Development with my child in the future **YES / NO** (circle one)

If at any time you have questions about the proposed research, please contact the study's Principal Investigator

Dr. Krista Byers-Heinlein
Centre for Research in Human Development
Department of Psychology, Concordia University
514-848-2424 x. 2208

If at any time you have questions about your rights as a research participant, please contact the Research Ethics and Compliance Advisor, Concordia University, 514.848.2424 ex. 7481 ethics@alcor.concordia.ca

Baby ID: ____________ Researcher: _________________________

Appendix B: Language Exposure Questionnaire

Language Exposure Questionnaire for Infants

A structured interview (do not hand directly to parents)

Baby_ID: _________________ Today's date (MM/DD/YYYY):_____________________________

Study_ID: ________________

Parent completing questionnaire: _________________________

Study Name: ______________

Family language background: Now I'm going to ask you some questions to get a better idea of [baby's name]'s exposure to different languages. First I'd like to ask about the languages spoken by people who spend time with [baby's name]. [Circle or write-in parents' answer. Put an X on box if not applicable, e.g., if parent does not use language in everyday life, put X on follow-up questions. If both caregivers not present, ask the one who is present to answer questions about other caregiver as accurately as possible.]

MOTHER (Caregiver 1)				
	English	**French**	**L3** ________________	**L4** ________________
Do you use [language] in everyday life?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never
What variety of [language] do you speak? [e.g. British English, Quebec French]				
Do you speak [language] to [child's name]?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never
At what age did you start learning [language]? [Enter 0 if native language/from birth.]				
When people hear you speak [language] can they guess that you speak another language?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never
FATHER (Caregiver 2)				
Do you use [language] in everyday life?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never
What variety of [language] do you speak? [e.g. British English, Quebec French]				
Do you speak [language] to [child's name]?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never
At what age did you start learning [language]? [Enter 0 if native language/from birth]				
When people hear you speak [language] can they guess that you speak another language?	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never	Regularly Sometimes Never

OTHER CAREGIVERS: Does [child's name] spend an hour or week or more on a regular basis with anyone else? What languages do they speak to him/her? [Note details]

____

DAYCARE: Does [child's name] attend regular childcare, such as daycare? **YES NO**

If yes, since what age: _________

What language(s) are spoken? [Note details]

____

TRIPS: Has [child's name] ever been on a trip of 1 month or more where his/her language exposure would have changed? **YES N**

Day-in-the-life estimate

Time wake up: ______ Total hours nap: ______ Bedtime: ______ **TOTAL WAKING HOURS/DAY:** ______ [generally ~ 12h but varies]

I want you to think about a day in **[child's name]**'s life, when s/he wakes up, who she's with, what s/he's doing, to get an idea of how many hours per day s/he hears **[language 1]** and how many hours per day s/he hears **[language 2]**. We're interested in people speaking directly to him/her, and <u>not in radio/TV</u>.

Let's start with when **[child's name]** was born. Were you at home with him/her? So on the weekdays, if s/he was awake, **X** hours, how many of those would you say were in **[language 1]** and how many in **[language 2]**? What about the weekends? [Walk parent through the baby's day to estimate hours per day in each language]

[Continue through life through increments of 1 month or more, starting a new row every time baby's situation changed (e.g., starting daycare, major change in family routine, long trip abroad). Continue until child's current age, rounding to the nearest month, where 15 days rounds down, and 16 days rounds up. Smallest unit of estimate is .5 hour.]

	E = English	Use "b" for birth		**Check:** Hours/day in each lang. sums to total waking hours/day								**Calculate:** hrs/week * #months			
Situation	**Languages**	**Ages**	**#months**	**M**	**T**	**W**	**Th**	**F**	**Sa**	**Su**		**L1** ___	**L2** ___	**L3** ___	**L4** ___
e.g. Home/began daycare/trip	E/F	b-12	12	4/8 →					6/6 →			384 (4h*5d +6h*2d) *12mos	624 (8h*5d + 6h*2d) * 12mos		

Global estimate: If you could put a tape recorder up to your baby's ear and counted all the words he/she has heard in her whole life in **[language 1]** and **[language 2],** what percentage do you think would be in each language?

____________ ____________ ____________ ____________

	L1	L2	L3	L4
Total hours in each language (sum each column)				**Check:** Sums to waking hrs/week * age _______
% Cumulative exposure (Calculate from total hours in each language)				**Check:** Sums to 100% _______
% Current exposure				**Calculate** based only on last chunk

Overall estimate (Average of <u>Global estimate</u> and day-in-the-life <u>% Cumulative exposure</u>).

* **Monolingual exception**: If % cumulative exposure >90%, use % cumulative without averaging.

Language:				
Percent:				
Age Acquisition:				

Appendix C: Language Mixing Questionnaire

Baby ID: ________________ Exp. Name: ________________________________
Study ID: ________________ Study Name: ________________________________

Caregiver's relation to infant (e.g. mother, father, grandmother): ____________________
Infant's date of birth :______________
Today's Date:_________________________

Language Mixing Questionnaire

a) In what situations do you tend to speak in English with your child? (check all that apply)
   ___ When one on one
   ___ At home
   ___ With friends
   ___ With family
   ___ At playgroup/lessons
   ___ When out (shopping, etc.)
   ___ Other (please specify)____________________________________

b) In what situations do you tend to speak in French with your child? (check all that apply)
   ___ When one on one
   ___ At home
   ___ With friends
   ___ With family
   ___ At playgroup/lessons
   ___ When out (shopping, etc.)
   ___ Other (please specify)____________________________________

c) What percentage of your interactions with your child are:
   in English? _____ %
   in French? _____%

Please answer the following questions, considering how you speak when interacting with your child. Please circle a number to indicate how much you agree with each statement.

d) I often start a sentence in English and then switch to speaking French.
   1 2 3 4 5 6 7
   Very true Somewhat true Not at all true

e) I often start a sentence in French and then switch to speaking English.
   1 2 3 4 5 6 7
   Very true Somewhat true Not at all true

f) I often borrow a French word when speaking English.

1 2 3 4 5 6 7
Very true Somewhat true Not at all true

I do this in situations when (check all that apply):
___ I'm not sure of the English word
___ No translation or only a poor translation exists for the word
___ The English word is hard to pronounce
___ When I'm teaching new words
___ Other times/not sure

g) I often borrow an <u>English</u> word when speaking <u>French.</u>
1 2 3 4 5 6 7
Very true Somewhat true Not at all true

I do this in situations when (check all that apply):
___ I'm not sure of the French word
___ No translation or only a poor translation exists for the word
___ The French word is hard to pronounce
___ When I'm teaching new words
___ Other times/not sure

h) In general, I often mix <u>English</u> and <u>French</u>.
1 2 3 4 5 6 7
Very true Somewhat true Not at all true

Appendix D: Demographic Questionnaire

Concordia Infant Research Laboratory Participant Information

Child's First Name: ________________________

First only

Child's Date of Birth: ____________________

MM / DD / YY

Child's Gender: ☐ Male ☐ Female ☐ Other/Not specified

Basic Family Information

Parent A's First Name: ___________________ ☐ Male ☐ Female ☐ Other/Not specified First only

Parent B's First Name: ___________________ ☐ Male ☐ Female ☐ Other/Not specified

First only

Address (including **postal code**):

____

____

Phone numbers	**Where?** (e.g. home, Mom work, Dad cell)
1.	
2.	
3.	
4.	
5.	

E-mail: ____

Does your child have any siblings?

First Name of Sibling	**Date of Birth**	**Gender**	**Can we contact you for future studies for this child?**
			☐Yes ☐No
			☐Yes ☐No
			☐Yes ☐No

<u>Languages Spoken in the Home and at Childcare</u>

What is parent A's **native language** (s)? _______________________

What is parent B's **native language** (s)? _______________________

What **percent of the time** does your baby **hear** the following languages?:

English		____%
French		____%
Other (please specify) ___________		____%
Other (please specify) ___________		____%
	TOTAL	**100 %**

Has the child lived/vacationed in any country **where s/he would hear a language other than English or French**? ☐**Yes** ☐ **No**

If yes, please detail (when, where, and for how long?) __________________________________

<u>Health History</u>

What was your **child's birth weight? __ __ lbs __ __ oz** OR **__ __ __ __ grams**

Was your child born early? ☐**Yes** ☐ **No**

If yes, how many days/ weeks? ____

Were there any **complications** during the pregnancy? ☐**Yes** ☐ **No**

If yes please detail ____

Has your child had any major **medical problems**?

If yes please detail____

Does your child have any **hearing, vision problems**, or **developmental delays**?

If yes please detail____

Does your child **<u>currently</u>** have an ear infection? ☐**Yes** ☐ **No**

Has your child had any ear infections **<u>in the past</u>**? ☐**Yes** ☐ **No**

If yes at which ages____

Does your child have a **cold** today? ☐**Yes** ☐ **No**

If yes, does he/she have pressure/pain in ears (if known)? ☐**Yes** ☐ **No**

Is there any other relevant information we should know (health or language-related)?

____

Has another university contacted you to participate in one of their studies? ☐**Yes** ☐ **No**

If yes, which university? __________________________________

Think of this ladder as representing where people stand in Canada.
At the **top** of the ladder are the people who are best off – those who have the most money, the most education, and the most respected jobs. At the **bottom** are the people who are the worst off – who have the least money, least education, and the least respected jobs or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the very bottom.

Where would you place yourself on this ladder?
Place a large "**X**" on the rung where you think you stand at this time in your life, relative to other people in Canada.

Family and Child Background Information (optional)

Parent A's Current Level of Education
Check any/all that apply:

- ☐ Primary School
- ☐ Some High School
- ☐ High School
- ☐ Some College/University
- ☐ College Certificate/Diploma
- ☐ Trade School Diploma
- ☐ Bachelor’s Degree
- ☐ Master’s Degree
- ☐ Doctoral Degree
- ☐ Professional Degree
- ☐ Not Applicable/Unknown
- ☐ Other (please specify): ______________________________ ______________________________ ______

Parent B’s Current Level of Education
Check any/all that apply:

- ☐ Primary School
- ☐ Some High School
- ☐ High School
- ☐ Some College/University
- ☐ College Certificate/Diploma
- ☐ Trade School Diploma
- ☐ Bachelor’s Degree
- ☐ Master’s Degree
- ☐ Doctoral Degree
- ☐ Professional Degree
- ☐ Not Applicable/Unknown
- ☐ Other (please specify): ______________________________ ______________________________ ______

Parent A's Occupational Status (optional)
Check any/all that apply:

- ☐ Employed Full-Time
- ☐ Employed Part-Time
- ☐ Stay-at-Home-Parent
- ☐ Student
- ☐ Unemployed
- ☐ Not Applicable/Unknown
- ☐ On Temporary Leave (e.g., maternity, paternity, sick, etc.; **please also check status when *not* on leave**)
- ☐ Other (please specify): ______________________________ ______________________________ ______

Parent B’s Occupational Status (optional)
Check any/all that apply:

- ☐ Employed Full-Time
- ☐ Employed Part-Time
- ☐ Stay-at-Home-Parent
- ☐ Student
- ☐ Unemployed
- ☐ Not Applicable/Unknown
- ☐ On Temporary Leave (e.g., maternity, paternity, sick, etc.; **please also check status when *not* on leave**)
- ☐ Other (please specify): ______________________________ ______________________________ ______

What language community do you (and your partner) identify with?
Check any/all that apply:

- ☐ Anglophone
- ☐ Francophone
- ☐ Allophone
- ☐ Other (please specify):

____

What are your child's ethnic origins?
Check any/all that apply:

- ☐ Aboriginal
- ☐ African
- ☐ Arab
- ☐ West Asian
- ☐ South Asian
- ☐ East and Southeast Asian
- ☐ Caribbean
- ☐ European
- ☐ Latin/Central/South American
- ☐ Pacific Islands
- ☐ Not Applicable/Unknown
- ☐ Other (please specify):

____

What culture do you (and your partner) identify with?
Check any/all that apply:

- ☐ Aboriginal
- ☐ African
- ☐ Arab
- ☐ West Asian
- ☐ South Asian
- ☐ East and Southeast Asian
- ☐ Caribbean
- ☐ European
- ☐ Latin/Central/South American
- ☐ Pacific Islands
- ☐ Not Applicable/Unknown
- ☐ Other (please specify):

____

Appendix E: Supplemental Analyses for Chapter 2

Stimuli used in Studies 1 and 2

Table E1 Full List of Stimuli Used in Studies 1 and 2

Montréal Word Pairs						New Jersey Word Pairs					
French			English			Spanish			English		
chien [m]	-	lait [m]	dog	-	milk	perro [m]	-	globo [m]	doggy	-	balloon
chat [m]	-	livre [m]	kitty	-	book	gato [m]	-	pie [m]	kitty	-	foot
bouche [f]	-	porte [f]	mouth	-	door	boca [f]	-	leche [f]	mouth	-	milk
biscuit [m]	-	pied [m]	cookie	-	foot	galleta [f]	-	puerta [f]	cookie	-	door

Note. The same word pairs were used in Study 1 and 2, with words in Study 1 presented in a list of words, and in Study 2 embedded in sentences. Slightly different items were used for the French-English and Spanish-English stimuli as necessary (e.g., to match grammatical gender or number of syllables across conditions). The table shows single-language trials, switched-language trials were created by substituting an English word with a French or Spanish word, e.g., substituting milk with *lait* [fr. milk] to form dog - *lait*. Stimuli for Study 2 were formed by inserting the target words in sentence frames: Do you like the ___? I want the ___! (English); *Aimes-tu le* ___? *Je veux le/la* ___! (French); ¿Te gusta el/la ____? Quiero el/la ___. (Spanish)

Differences in Samples between Studies 1 and 2

The samples for Studies 1 and 2 were largely overlapping, with some small differences. Eight infants whose data were included in Study 1 were excluded from Study 2 for contributing fewer than 8 trials that met our minimum threshold of 2 s of looking time. The final sample for Study 2 included data from four infants whose data had been excluded in Study 1 (for contributing too few trials, $n = 2$, or due to technical difficulties that only affected Study 1, $n = 2$). The dataset posted on OSF contains the full information about the samples in both studies (at https://osf.io/9dtwn/).

ANOVA for infants tested in Montréal only

Initially, we planned to compare infants tested in Montréal and New Jersey. However, we were unable to recruit a sufficient number of infants in New Jersey who met the bilingual inclusion criteria. To test whether combining a small number of English–Spanish bilinguals with our full sample of English–French bilingual infants obscured any patterns of results present when analyzing only infants tested in Montréal, we conducted an ANOVA that included only infants tested in Montréal. Excluding infants tested in New Jersey yielded the same pattern of results reported in Study 1 (trial type: $F[1, 32] = 0.17$, $MSE = 0.04$, $p = .684$, language background: $F[1, 32] = 3.55$, $MSE = 0.09$, $p = .069$, language background × trial type interaction: $F[1, 32] = 0.01$, $MSE = 0.04$, $p = .931$). The marginal effect of language background reflects that bilinguals had slightly longer looking times than monolingual infants ($M_{bilingual} = 9.53$, $M_{monolingual} = 8.36$); however this difference did not vary by trial type. In the ANOVA for Study 2 excluding infants tested in New Jersey, there was no significant main effect of trial type ($F[1, 31] = 1.24$, $MSE = 0.04$, $p = .275$), and no other significant effects (language background: $F[1, 31] = 1.66$, $MSE = 0.16$, $p = .207$, language background × trial type: $F[1, 31] = 0.00$, $MSE = 0.04$, $p = .997$). These results are consistent with the results from the full sample, and indicate that adding data from infants tested in New Jersey did not change the pattern of results present in infants tested in Montréal.

Descriptive data for sample collected in New Jersey

We were unable to complete the full sample of English–Spanish bilingual infants tested in New Jersey due to challenges in recruitment. Due to the small sample, we are unable to calculate statistical tests on this sample, but here we report descriptive data. In Study 1, after excluding participants for not meeting pre-specified health and language we were left with 8 participants. These participants completed on average 15 trials (range: 10–16) prior to the exclusion of trials that did not satisfy the minimum trial length, and 7 participants who contributed 15.1 trials (range: 10–16) after removal. Averaging looking times across infants for the two trial types in Study 1, infants tested in New Jersey looked numerically but not statistically longer to switch than single trials, see Table E2, a trend that is consistent with the results from Montréal.

In Study 2, after excluding participants for health, language and technical reasons, we were left with 8. These participants completed on average 12 trials (range: 5–16) prior to the exclusion of trials that did not satisfy the minimum trial length, and 4 who completed 13.8 trials (range: 9–16) after removal. The average looking times for each trial type in Study 2 again showed a numerical trend of longer looking for switched-language than single-language trials, see Table E2. The trends observed in the English–Spanish bilingual sample tested in New Jersey in Studies 1 and 2 are consistent with those observed in the sample tested in Montréal.

Table E2 Looking Times to Switch and Single-Language Trials by Test Location

		Looking Time (s)			
		Single-Language		Switched-Language	
Test Location	N	Mean	SD	Mean	SD
		Study 1			
Montréal	34	8.80	2.76	8.89	1.94
New Jersey	7	7.34	1.18	7.72	1.86
		Study 2			
Montréal	33	8.09	2.84	8.70	2.45
New Jersey	4	5.38	0.84	6.41	1.26

Appendix F: Supplemental Analyses for Chapter 3

Information about incomplete samples

For some groups in Study 1, we had to stop data collection due to recruitment difficulties. These participants included 3 9-month-olds (all bilinguals), 37 18-month-olds (9 bilingual, 17 monolingual, 11 who did not fit the criteria for either bi- or monolingual), as well as bilinguals with English or French and another language: 15 bilingual-other 5-months-olds and 9 bilingual-other 12-month-olds.

Robustness analyses

The following tables contain the same analyses as in the results section, but with untransformed looking time data instead of the log-transformed looking time data used in the results section.

Table F1 Robustness Analysis for ANOVAs on Looking Time

Variable	Result
	ANOVA for Study 1, 5-month-olds
Trial Type	$F(1, 36) = 0.00$, $MSE = 4.07$, $p = .995$, $\eta^2_p = .000$
Language Background	$F(1, 36) = 0.56$, $MSE = 17.55$, $p = .459$, $\eta^2_p = .015$
Language Background × Trial Type	$F(1, 36) = 0.01$, $MSE = 4.07$, $p = .936$, $\eta^2_p = .000$
	ANOVA for Study 1, 12-month-olds
Trial Type	$F(1, 44) = 0.46$, $MSE = 3.11$, $p = .502$, $\eta^2_p = .010$
Language Background	$F(1, 44) = 0.49$, $MSE = 8.67$, $p = .486$, $\eta^2_p = .011$
Language Background × Trial Type	$F(1, 44) = 1.12$, $MSE = 3.11$, $p = .295$, $\eta^2_p = .025$
	ANOVA for Study 2, 12-month-olds
Trial Type	$F(1, 40) = 2.47$, MSE = 4.90, $p = .124$, $\eta^2_p = .058$
Language Background	$F(1, 40) = 3.12$, $MSE = 19.08$, $p = .085$, $\eta^2_p = .072$
Language Background × Trial Type	$F(1, 40) = 0.19$, $MSE = 4.90$, $p = .667$, $\eta^2_p = .005$
	ANOVA for Study 2, 18-month-olds
Trial Type	$F(1, 37) = 1.40$, $MSE = 9.18$, $p = .244$, $\eta^2_p = .036$
Language Background	$F(1, 37) = 1.72$, $MSE = 39.06$, $p = .197$, $\eta^2_p = .044$
Language Background × Trial Type	$F(1, 37) = 0.85$, $MSE = 9.18$, $p = .363$, $\eta^2_p = .022$

Table F2 Results of Robustness Analysis for Equivalence Tests on Untransformed Data

Study	Age Group (months)	*t*-Test
1	5	$t(37) = 3.15$, $p = 0.002$

Study	Age Group (months)	*t*-Test
	12	*t*(45) = -2.80, *p* = 0.004
2	12	*t*(41) = -1.73, *p* = 0.046
	18	*t*(38) = 2.01, *p* = 0.026

Table F3: Results of Robustness Analysis for Correlation Between Exposure to Dominant Language and Differential Looking to Switch and Same Trials

Study	Age Group (months)	*r*	*95% CI*	*t*-Test
1	5	-.3	[-.35, .29]	*t*(36) = -0.18, *p* = .857
	12	.19	[-.10, .46]	*t*(44) = 1.29, *p* = .203
2	12	.2	[-.28, .33]	*t*(40) = 0.16, *p* = .876
	18	-.16	[-.45, .16]	*t*(37) = -1.00, *p* = .322

Table F4 Results of Robustness Analysis for Correlation Between Exposure to Dominant Language and Differential Looking to Switch and Same Trials

Study	Age Group (months)	*r*	*95% CI*	*t*-Test
1	5	.4	[-.44, .49]	*t*(16) = 0.15, *p* = .886
	12	.27	[-.20, .63]	*t*(18) = 1.17, *p* = .256
2	12	.8	[-.40, .53]	*t*(16) = 0.33, *p* = .748
	18	.1	[-.43, .45]	*t*(18) = 0.06, *p* = .950

Table F5 Robustness Analysis for ANOVA on Direction of Switch and Language Background

Variable	Result
	ANOVA for Study 1, 5-month-olds
Direction of Switch	*F*(1, 36) = 1.35, *MSE* = 13.58, *p* = .253, η^2_p = .036
Language Background	*F*(1, 36) = 0.01, *MSE* = 16.28, *p* = .936, η^2_p = .000
Direction of Switch × Language Background	*F*(1, 36) = 3.39, *MSE* = 13.58, *p* = .074, η^2_p = .086
	ANOVA for Study 1, 12-month-olds
Direction of Switch	*F*(1, 42) = 0.01, *MSE* = 11.95, *p* = .915, η^2_p = .000
Language Background	*F*(1, 42) = 1.31, *MSE* = 12.90, *p* = .258, η^2_p = .030
Direction of Switch × Language Background	*F*(1, 42) = 0.15, *MSE* = 11.95, *p* = .699, η^2_p = .004
	ANOVA for Study 2, 12-month-olds
Direction of Switch	*F*(1, 30) = 0.00, *MSE* = 31.02, *p* = .965, η^2_p = .000
Language Background	*F*(1, 30) = 0.18, *MSE* = 22.22, *p* = .675, η^2_p = .006
Direction of Switch × Language Background	*F*(1, 30) = 0.41, *MSE* = 31.02, *p* = .526, η^2_p = .014
	ANOVA for Study 2, 18-month-olds
Direction of Switch	F(1, 27) = 0.00, MSE = 23.45, *p* = .991, η^2_p = .000
Language Background	F(1, 27) = 0.05, MSE = 30.57, *p* = .832, η^2_p = .002
Direction of Switch × Language Background	F(1, 27) = 4.29, MSE = 23.45, *p* = .048, η^2_p = .137