Untangling Bilingualism: Using Code-Switching to Understand Bilingual Language Development

Lena van der Velde Kremin

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

April 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Lena Kremin

Entitled: Untangling Bilingualism: Using Code-Switching to Understand Bilingual Language Development

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. Ian Ferguson

____ External Examiner
Dr. Natalie Brito

____ Thesis Supervisor
Dr. Krista Byers-Heinlein

____ Examiner
Dr. Pavel Trofimovich

____ Examiner
Dr. Mickael Deroche

____ Examiner
Dr. Kristen Dunfield

Approved by ____

Dr. Andrew Chapman, Graduate Program Director

6/9/2022 ____

Dr. Pascale Sicotte, Dean

Faculty of Arts and Science

Abstract

Untangling Bilingualism: Using Code-Switching to Understand Bilingual Language Development

Lena van der Velde Kremin, Ph.D.
Concordia University, 2022

Switching between two languages, or code-switching, is common in bilingual communities. However, little is known about the code-switching young bilinguals hear in their daily lives and how they process it. This dissertation investigated these two aspects of code-switching and proposed new models for defining bilingualism.

Bilingualism is difficult to define and model. In Chapter 2, I proposed that bilingualism researchers can integrate psychometric models, such as the factor mixture model and the grade-of-membership model, which incorporate both categorical and continuous properties. Such models can unify traditional approaches of defining bilingual groups with newer views of bilingualism as a continuous variable. These models will allow researchers to address a variety of research questions, advance theory, and lead to a deeper understanding of bilingualism.

In Chapter 3, I analyzed French–English parents' code-switching in day-long at-home audio recordings, provided when their infant was 10 and 18 months old. Code-switching was relatively infrequent: an average of 7 times per hour (6 times/1,000 words) at 10 months, increasing to 28 times per hour (18 times/1,000 words) at 18 months. Parents code-switched more between sentences than within a sentence, and this pattern became more pronounced when infants were 18 months. Parents appeared to code-switch most frequently to bolster their infant's understanding and teach vocabulary, suggesting that code-switching may support successful bilingual language development.

In Chapter 4, I investigated how bilingual children process code-switching, examining how 3-year-old bilinguals process sentences with code-switches at an uninformative determiner-adjective pair before the target noun (e.g., "Can you see *el buen* [sp. the good] duck?) compared to single-language sentences (e.g., "Can you see the good duck?"). Children were unexpectedly accurate at identifying the target noun in both sentence types, contrasting with previous findings that code-switching leads to processing difficulties. Surprisingly, exploratory results suggested that code-switching may have boosted comprehension for certain children.

In sum, this dissertation has illustrated how code-switching may support bilinguals' language development. I discovered that parents code-switch to support their child's learning and showed that children do not always have difficulty processing code-switching. Bilingualism is a multi-faceted phenomenon, and nuanced research is needed to capture this variability.

Acknowledgements

First and foremost, I would like to give the biggest thanks to my supervisor, Dr. Krista Byers-Heinlein, for guiding me through the last four years. From the first time we spoke, your passion for your research was clear, infectious, and somewhat intimidating in the best possible way. You not only cared that I was doing good work but gently pushed me to grow as a researcher, always meeting me where I was and showing me the path to improvement. I've always felt that you wanted what was best for me, whether that was in research productivity, career shifts, or a healthy work-life balance. In that sense, your mentorship went far beyond conducting research to teaching me how to be a better mentor myself. This philosophy has made its way into the team you have built, a group of highly intelligent and motivated researchers who truly want everyone to succeed. Your supervision and this environment made the last four years such a rewarding experience. I often tell people that I'm lucky to have had such a great supervisor, but it's not entirely luck. You work hard to be there for your students, and for that I will be forever grateful.

An overarching theme of my time in grad school has been the importance of doing good, collaborative science. Two chapters in this dissertation would not have been possible without the contributions from others. Many thanks to Amel Jardak, and Drs. Casey Lew-Williams, AJ Orena, and Linda Polka for your insight, feedback, and even data that made these chapters possible. I also thank my internal committee members, Drs. Mikael Deroche and Kristen Dunfield, for keeping their minds open when the pandemic shifted my research plans and for offering suggestions for how to move forward.

I owe a lot of my growth and laughs to my lab mates: Melanie Brouillard, Hilary Killam, Melanie López Pérez, Andrea Sander Montant, Esther Schott, and Drs. Nicolás Alessandroni, Rodrigo Dal Ben, Laia Fibla Reixachs, Ana Maria Gonzalez Barrero, Charlotte Moore, Rachel K.-Y. Tsui, and Erin Quirk. It has been so wonderful to work in such a supportive, collaborative, and friendly lab. While the time I've known each of you has varied, I've enjoyed getting to know all of you and have learned something from everyone. I'll miss all the free lunch runs, ice cream trips, and puns.

I am also grateful to have worked with two fantastic honor's students, Julia Alves and Erica Chelini. It was such a rewarding experience to see you grow as researchers, and I greatly appreciate your support on my projects and that you occasionally did the most finicky parts of the process.

While their support did not directly impact the work in this dissertation, I would also like to thank my past mentors and supervisors for providing me with the learning opportunities and letters of reference that led to this experience: Drs. Marlyse Baptiste, Nick Ellis, Sam Epstein, René Kager, Ioulia Kovelman, and Núria Sebastián-Gallés. I also acknowledge that this dissertation was funded by fellowships from the Fonds de Recherche du Québec and the Centre for Research on Brain, Language, and Music.

Lastly, to my friends and family, my mom, Marijke, Cebby, Sandra, Lindsey, and Tristan, thank you for keeping me relatively sane. I will forever appreciate the texts, calls, walks, and picnics where you lent an ear and/or shoulder as I navigated the crazy world of academia. And even though he can't read this, to Samson, my nugget, my croissant, my ball of fluff and cuddles, thanks for being my pal and constant supply of smiles through all of this.

Dedication

To my past, present, and future self.

Contribution of Authors

This dissertation consists of three manuscripts containing one theoretical paper and two empirical studies. Authors' relative contributions to each manuscript are noted.

Manuscript 1 (Chapter 2)

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

LVK: Conceptualization, Writing – Original Draft

KBH: Conceptualization, Writing – Review & Editing, Supervision

Manuscript 2 (Chapter 3)

Kremin, L. V., Alves, J., Orena, A. J., Polka, L., & Byers-Heinlein, K. (2021). Code-switching in parents' everyday speech to bilingual infants. *Journal of Child Language*, 1–27. https://doi.org/10.1017/S0305000921000118

LVK: Conceptualization, Methodology, Formal analysis, Writing – Original Draft, Visualization

JA: Data coding

AJO: Conceptualization, Writing – Review & Editing

LP: Conceptualization, Writing – Review & Editing, Project Administration, Funding Acquisition

KBH: Conceptualization, Writing – Review & Editing, Supervision, Project Administration, Funding Acquisition

Manuscript 3 (Chapter 4)

Kremin, L. V., Jardak, A., Lew-Williams, C., & Byers-Heinlein, K. (*under review*). Bilingual children's comprehension of code-switching at an uninformative adjective.

LVK: Formal analysis, Writing – Original draft, Visualization

AJ: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft

CLW: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition

KBH: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition

Table of Contents

List of Figures

List of Tables

1 General Introduction

With globalization and increasing migration, the number of people who speak two or more languages is on the rise. This also applies to children who are exposed to multiple languages from a young age. An estimated 1 out of 5 children in Canada (Schott, Kremin, et al., 2021) and the US (Kids Count Data Center, 2018) are bilingual. One reason it is difficult to establish the exact rate of bilingualism is because definitions and measures of bilingualism are not consistent in the literature. Even when scholars use different definitions of bilingualism, relatively little research is focused on how young bilinguals learn both of their language (Kidd & Garcia, 2021). One particular feature of bilingual language development that is beginning to receive more attention is how the switches between languages, or code-switches, that a young bilingual hears affects their language development. This dissertation aims to better understand bilingualism, particularly through the lens of children's language development, and includes one methodological chapter and two empirical chapters to address these points. Chapter 2 proposes a novel approach to defining and measuring bilingualism. Chapter 3 investigates the code-switching that young bilinguals hear in their daily life, and Chapter 4 examines how they process such code-switching. In the next sections, I will review the literature on current approaches to defining bilingualism, elements of bilingual language development, including language input and comprehension, and code-switching.

1.1 Defining and Measuring Bilingualism

A commonly used definition of a bilingual is a person who uses two or more languages in their everyday life (Grosjean, 1989). When taking a closer look, however, there is great ambiguity and debate over what is meant by "use" a language. Is someone bilingual if they can understand but not speak a second language? Is someone bilingual if they can speak one language fluently and the second with some hesitation? Is someone bilingual if they are taking classes to learn a second language? If you were to ask someone on the street to define bilingualism, they might expect that a bilingual has perfect command over both of their languages. Indeed, this definition of "balanced bilingualism" remains widely held in the general population and was once the dominant definition in academic spaces as well (Lambert et al., 1959). Scholars have since dispelled the idea that a bilingual is "two monolinguals in one" (Grosjean, 1989) but have not converged on a single way to define bilingualism.

Researchers generally agree that bilingualism is a multidimensional construct, that bilingualism includes a combination of many skills and depends on many different factors (e.g., Antoniou, 2019; Luk & Bialystok, 2013). This multidimensional nature can be seen in how bilingualism is typically measured. For example, a young child's bilingual status can be assessed via comprehensive test batteries that measure a range of linguistic skills, such as listening and speaking skills (e.g., Woodcock et al., 2005), and an adult's bilingual status can be assessed through their language history, including when they began learning each language they speak and how often they speak each language (e.g., Li et al., 2014). While measuring multiple dimensions when determining bilingual status is common, if not expected, in the field, the exact dimensions that are used and how they are combined varies widely between studies. In some cases, this is understandable. For example, adults can answer detailed questions about their language history and complete multiple tasks that assess their proficiency in each of their languages (e.g., Language and Social Background Questionnaire, Anderson et al., 2018; Language History Questionnaire, Li et al., 2014; Language Experience and Proficiency Questionnaire, Marian et al., 2007). Infants, however, cannot complete the same measures, so their bilingual status is often evaluated through a structured interview with a caregiver to gather information on what languages the infant hears, how much they are exposed to them, and when this exposure began (Language Experience Questionnaire, Bosch & Sebastián-Gallés, 2001; Multilingual Approach to Parent Language Estimates, Byers-Heinlein et al., 2020). The challenge that researchers face is determining exactly what combination of skills and information they will use to evaluate bilingual status for the purposes of their study. This issue of the multidimensional nature of bilingualism is further complicated by the additional step of deciding how to define bilingualism in the study by using the measures that they chose.

Researchers have, in the past several decades, expanded the definition of bilingualism beyond "balanced bilingualism" to include a larger number of bilinguals. With this expansion comes increased variability in the bilinguals themselves. Researchers have tried to capture this variability by creating different groups of bilinguals. This is evidenced by the more than 100 different labels for bilinguals in the literature (Surrain & Luk, 2017). Labels such as "second language learner" and "simultaneous bilingual" can be useful to describe particular groups that are being studied, but researchers do not always agree on the thresholds for membership in these groups. For example, infants' bilingual status is often evaluated by calculating the percent of their entire language exposure they receive in each of their languages, as described above (Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., 2020). While many studies use a measure of exposure, the cutoff value used to place an infant in the monolingual or bilingual group has varied across studies from 10% - 40% exposure to the lesser-heard language (Rocha-Hidalgo & Barr, 2021). Inconsistent cutoffs like this have two main effects on research. First, they dichotomize samples into monolingual and bilingual groups, potentially resulting in incorrect conclusions, as an infant considered "bilingual" in one study could be considered "monolingual" in another (Rocha-Hidalgo & Barr, 2021). Second, they make it difficult to synthesize reported findings.

Although placing participants into discrete groups has long been an established practice in the field of bilingualism, categorical comparisons may not accurately represent the underlying structure of bilingualism. Recent proposals have called for bilingualism to be treated as a continuous variable to better capture to variability in and nature of the construct (Baum & Titone, 2014; de Bruin, 2019; Kaushanskaya & Prior, 2015; Marian & Hayakawa, 2021; Takahesu Tabori et al., 2018). Determining whether the underlying structure of bilingualism is categorical or continuous is an important question with statistical consequences. If bilingualism is underlyingly continuous, treating it as categorical can lead to reduced statistical power, smaller effect sizes, increased chances of a Type I error, and spurious interactions and main effects (MacCallum et al., 2002). Thus, approaching bilingualism as a continuous construct could address these statistical concerns and allow for subtle and threshold effects of bilingualism to be investigated (de Bruin, 2019; De Cat et al., 2018).

Regardless of whether researchers approach bilingualism as a categorical or continuous construct, they still face the challenge of comparing definitions of and results on bilingualism across research studies. Arriving at a somewhat consistent definition of bilingualism may be a key step in advancing the field, as some believe that inconsistent definitions play a large role in the inconsistent findings in the literature (de Bruin, 2019; Kaushanskaya & Prior, 2015; Marian & Hayakawa, 2021). In categorical approaches, inconsistent cutoffs make it difficult to compare results between different studies. Depending on what criteria is used, the same participant could be classified as a monolingual in one study but a bilingual in another. In continuous approaches, studies may use different measures or the same measures in different ways to arrive at their final continuum. To address the issue of inconsistency across studies, there have been several recent proposals to standardized the definition and measurement of bilingualism in the field (e.g., De Cat et al., 2021; Marian & Hayakawa, 2021). However, even these proposals cannot all agree on what standardization would look like.

The differences in the measures used to evaluate bilingual status and whether bilingualism is defined categorically or continuously highlight the need for new approaches in defining and measuring bilingualism. To address any research question about bilingualism, scholars must 1) decide how to define bilingualism, 2) select which measures to use, and 3) choose either a categorical or continuous model. Despite calls and efforts for a standardized measure in the field, a single definition, set of measures, and model may be impractical, and in some cases limiting, for all the different forms bilingualism research can take. Chapter 2 proposes that bilingualism researchers can draw from other fields to find a blended approach that will allow for both consistency within and flexibility across particular subfields.

1.1.1 Definitions of Bilingualism in the Current Work

As discussed above, bilingualism can be defined and modeled in many different ways across studies. Within a particular empirical study, however, a consistent definition of bilingualism that matches the aim of the study is important to ensure that the sample collected can properly address the research question. The main purpose of the two empirical studies included in this dissertation was to look at the effect of a bilingual experience, rather than to evaluate the effects of amount of exposure or level of proficiency. Thus, I opted to use a categorical approach to select participants who had a minimum amount of exposure to or a minimum level of proficiency in each of their languages. Bilingualism was defined as having at least 25% exposure to both languages from birth (Chapter 3) or having a parent-reported proficiency of at least 7/10 in both languages (Chapter 4). The different measures of bilingualism used in the two chapters reflect how the definition of bilingualism can and should change throughout development (as discussed in Chapter 2). In Chapter 3, the participants were first studied at 10 months and again at 18 months. These ages are too young for any ratings of proficiency, so language exposure was used to determine their bilingual status as it is the most readily available measure of their bilingual status (Byers-Heinlein, 2015). In Chapter 4, the participants were between 3 and 4 years old. At this age, children are able to put sentences together and hold basic conversations. Thus, a measure of their proficiency in each language, as rated by their parents, was used to determine their bilingual status (Marian et al., 2007). Furthermore, bilingualism was defined categorically in both chapters and no monolinguals were included in the final samples.

1.2 Bilingual Language Development

Even under the most liberal definition of bilingualism, two languages are involved, and bilingual infants and children are tasked with learning both. In the earliest stages of language development, input is a crucial component for successful learning: when it comes to language, children cannot learn what they do not hear or see. Input is such a vital element in the early stages, that the amount of input an infant receives in both languages is frequently how their bilingual status is determined (as mentioned above). Additionally, a general trend that appears consistently in the literature is that bilinguals have more developed language abilities and larger vocabulary sizes in the language they hear more (Côté et al., 2022; Hoff et al., 2012; Thordardottir, 2011). Thus, the language input of young bilinguals is an important element in studying and fully understanding bilingual language development. Because young bilinguals must learn from their input, it is also important to investigate how children process and ultimately comprehend the language input that they hear. Part of this question of language processing involves how young bilinguals recognize that they are learning two languages and how they come to represent them as separable entities. This section explores early bilingual language development and the importance of language input, language processing and comprehension, and language separation.

1.2.1 Language Input

One important element of input on language development is the amount of input infants receive. While infants can learn language from speech they overhear in their environment (Floor & Akhtar, 2006), the amount of input that is directed towards them appears to play a large role in their language development (Golinkoff et al., 2015; Shneidman & Goldin-Meadow, 2012). The amount of infant-directed speech an infant hears has been linked to later word production (Ramírez-Esparza et al., 2014), and infants who partake in more conversational turns are better at processing language (Romeo et al., 2018). Infants' language input varies widely beyond the overall amount of input they receive, including number of speakers and the larger social context, yet capturing and studying all these potential differences is extremely difficult (Hoff, 2020). Moreover, most of the research to date on language input has focused on monolingual infants, who hear only a single language.

Bilingual infants hear two languages, so researchers must not only consider the total amount of input these infants hear, but also the amount of input heard in each language. Relative exposure to each language can be challenging to measure, because it often relies on parental reports, which may be inaccurate if not collected carefully (Byers-Heinlein et al., 2020). Researchers have used methods such as surveys (Byers-Heinlein, 2013), diary studies (De Houwer, 2011; Place & Hoff, 2011), observation play session in the laboratory (Bail et al., 2015), and daylong audio recordings (Orena et al., 2020). Each of these methods has its own advantages and disadvantages, but all have been able to shed light on the language input that young bilinguals hear in their daily lives. For example, De Houwer (2011) used both surveys to reach a large number of bilingual families and a diary study to gather more detail about bilinguals' language exposure. Surveys from over 2,500 two-parent, bilingual families in Belgium measured patterns of family language exposure, classifying children based on, for example, whether they heard both languages from both parents or a different language from each parent. In the diary study, researchers instead measured the amount of time children heard each of their languages and the proportion they heard their languages across different days. Overall, the study found that differences in language input across children could explain differences in the children's language use. These results highlight the variability in different bilingual environments but do not offer extensive fine-grained detail on bilinguals' input in each of their languages.

A much more fine-grained level of detail was achieved in a recent study, where Orena and colleagues (2020) used daylong audio recordings to capture the language input that bilinguals heard in their home, including the proportion of their input they heard in each of their languages. Twenty-one French–English bilingual families recorded the language environment of their 10-month-old infant for 16 hours a day for three days. Trained research assistants then hand-coded these recordings to determine the amount of input the bilinguals heard in each of their languages. One significant finding from this study was that the input that young bilinguals heard varied greatly depending on what type of input was being measured. For example, the proportion of their input that a child heard in each of their languages varied between infant-directed speech and overheard speech. Additionally, the proportion heard in each of their languages varied day to day depending on who was taking care of the child. These results highlight that bilingual language input is a dynamic and complex factor to consider when investigating bilingual language development. Importantly, this study also confirmed that caregivers were fairly accurate at reporting the proportion of time their child hears both of their languages, meaning that this information can be easily collected through a structured interview (Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., 2020). Understanding bilinguals' exposure to each of their languages allows researchers to study how this input affects their overall language development.

Researchers are now beginning to move beyond macro-level measures of bilingual infants' language environments, such as the amount of input in each language, to investigate the role of more micro-level features to develop a more detailed understanding on the effect of input on language development. One such feature that has been investigated is code-switching. Bilinguals inevitably hear switches between their two languages – whether the switch happens across larger time scales (e.g., when moving from home to school) or smaller time scales (e.g., within a single conversation or sentence). Several studies have investigated the impact of exposure to code-switching on bilinguals' language development, focusing mainly on vocabulary development, and have found mixed results. One study found that higher rates of code-switching were linked to smaller comprehension vocabularies at 18 months and smaller productive vocabularies at 24 months (Byers-Heinlein, 2013). However, other studies did not find any relationship between the amount of code-switching a child hears and their vocabulary size (Bail et al., 2015; Carbajal & Peperkamp, 2020; Place & Hoff, 2016). The contrasting findings in these studies may be due to different methodologies. Parental surveys rely on parents' ability to reflect on their own speech patterns and report them honestly, which could lead to errors in their reporting (Byers-Heinlein, 2013). Laboratory play sessions have the benefit of direct observation (Bail et al., 2015), but the speech children hear during structured play is different than what they hear throughout the remainder of the day (Bergelson et al., 2019; Tamis-LeMonda et al., 2017). Thus, a new approach to study the code-switching that young bilinguals hear in their daily life is needed to better understand the language input they are receiving and any potential impacts on their language development. Chapter 3 addresses this issue by developing a method to analyze parents' code-switching in daylong recordings of infants' language environments.

1.2.2 Language Processing and Comprehension

Children extract patterns from their language input and come to understand the meaning of the words that they hear. One of the earliest indications of processing and subsequent comprehension is that infants are able to understand common nouns by 6- to 9-months (Bergelson & Swingley, 2012, 2015).

A common way to assess children's language processing and comprehension is with the looking-while-listening procedure. In this procedure, children are placed in front of a screen, shown pictures of objects, and played corresponding words over a speaker. Then, via a video recording or automatic eye tracker, children's gaze is recorded and analyzed (Swingley, 2012). There are several different approaches to analyzing the gaze data, but the general approach is to evaluate if children spend more time looking at the object that was labeled than the object that was not. For example, a child sees the images of a cat and a dog and hear the sentence, "Look at the dog!" After hearing the sentence, the child should look at the image of the dog longer than the image of the cat if they understand the word "dog." The looking-while-listening procedure has been used to investigate children's comprehension of different aspects of language, such as nouns (Fernald, McRoberts, et al., 2001), verbs (Golinkoff et al., 2001), and adjectives (Fernald et al., 2010).

A number of studies have used the looking-while-listening procedure to investigate aspects of bilingual children's language processing, for example, their comprehension in each of their languages, and whether this is linked to other elements of their language development. To illustrate, for Spanish–English children, higher amounts of exposure to a language have been shown to be related to better comprehension in that language (Marchman et al., 2017). Similarly, as vocabulary in one language grows, Spanish–English children's comprehension of sentences in that language increases as well (Marchman et al., 2010). As discussed above, bilingual children not only hear sentences in a single language, but they also hear sentences with code-switching. Thus, researchers have recently begun to using the looking-while-listening procedure to investigate how young bilinguals process and understand speech that contains code-switching, mainly code-switching at a noun (Byers-Heinlein et al., 2017; Byers-Heinlein, Jardak, et al., 2021; Morini & Newman, 2019; Potter et al., 2019). Chapter 4 extends this literature by investigating how bilinguals process code-switching at previously uninvestigated location, a prenominal determiner-adjective pair.

1.2.3 Language Separation

For bilingual infants, recognizing that they are hearing and learning two languages is a key step in their language development (Byers-Heinlein, 2014), but the exact nature and timing of this separation is not entirely understood. Early theories on adult bilinguals' language representations posit that bilinguals have two separate, encapsulated language systems (Macnamara, 1967; Penfield & Roberts, 1959). Under these theories, the encapsulated languages turn on and off as bilinguals switch between their languages. However, research has since shown that bilinguals' languages are connected and interact across all levels of language (Kroll et al., 2012). Regardless of exactly how languages are represented, it is clear that bilinguals are able to treat them as functionally separate categories (Byers-Heinlein, 2014), drawing the sounds (Burns et al., 2007; Gonzales et al., 2019), words (Genesee et al., 1995), and grammar from one language or another as appropriate (Grosjean, 2001), even while they might also produce code-switches and other productions that involve blending the two languages (Ritchie & Bhatia, 2012). In the context of language development, the question then becomes how young bilinguals recognize that they are learning two languages and come to represent them as functionally separate categories.

Most current theories of separation in language development propose that bilingual infants differentiate their languages from early in development (Genesee, 1989). Support for these theories of early separation comes from infants' ability to discriminate their languages based on rhythmic patterns within the first days and months of life (Bosch & Sebastián-Gallés, 1997b, 2001; Byers-Heinlein et al., 2010). However, the ability to discriminate linguistic input does not necessarily imply that bilingual infants recognize it as coming from two different languages, as rhythmic discrimination taps into infants' innate perceptual sensitivities (Byers-Heinlein, 2014; Nazzi et al., 1998). Instead of being taken as evidence that they have functionally separate categories for their languages, bilingual infants' ability to discriminate their languages based on rhythm may support the subsequent creation of these categories as they acquire the phonology, syntax, and lexicon of each language (Curtin et al., 2011). Thus, researchers can investigate when bilinguals develop these categories by assessing if young bilinguals treat stimuli from their languages differently as bilingual adults do (Byers-Heinlein, 2014).

One way of addressing this question is by looking at how young bilinguals handle code-switching. Because code-switching involves both languages, it provides the opportunity to see how bilinguals respond when they must switch between their two language categories. Investigating responses to code-switching throughout development could shed light on the timing and nature of the emergence of their language categories.

1.3 Code-Switching

Code-switching is common in bilingual and multilingual environments (Cheng & Butler, 1989), and it occurs regularly in the language input that young bilinguals receive (Bail et al., 2015; Byers-Heinlein, 2013; David & Wei, 2008; Goodz, 1989). Code-switching used to be viewed as evidence of weakened linguistics systems, or as a sign of lack of proficiency in one or both languages (Heredia & Altarriba, 2001; Nilep, 2006; Weinreich, 2010). Today, code-switching is viewed as a linguistic tool that bilinguals can use to add additional meaning and/or nuance to their speech (Blom & Gumperz, 1972; Gumperz, 1982; Myers-Scotton, 2017; Nilep, 2006). Moreover, code-switching is now widely regarded as a highly systematic and complex linguistic phenomenon (Poplack, 1980; Ritchie & Bhatia, 2012). This dissertation investigates the impact of code-switching on young bilinguals' language development from two related viewpoints: the code-switching input young bilinguals receive (investigated in Chapter 3) and how they process and understand it (investigated in Chapter 4).

1.3.1 Code-Switching Production

When investigating the code-switching that a young bilingual hears from the people in their environment, it is important to understand the different factors and features present in the production of code-switching. One very salient factor is the frequency with which code-switching occurs. Bilinguals produce code-switching at different rates, with some speakers code-switching very frequently, but others code-switching only rarely. This variation is present in speech that is directed towards both adults and children (Bail et al., 2015; Dewaele & Li, 2014; Dewaele & Zeckel, 2016). As mentioned above, the frequency that a young bilingual hears code-switching may affect their language development (Byers-Heinlein, 2013), so investigating the frequency that code-switching is produced in their environment is important to understanding the process of bilingual language development.

Beyond variation in the frequency of code-switching, code-switching can be produced at many syntactic locations. It can occur both between sentences (i.e., intersententially) and within a single sentence (i.e., intrasententially). Code-switches that occur intersententially are not subject to syntactic constraints, as the sentence before the code-switch and the one after can follow the grammars of their respective languages (Azuma, 1992, 2009; Myers-Scotton, 1997). The grammar of code-switching becomes more complex when the code-switch happens intrasententially. In this case, the code-switch must accommodate the grammars of both languages (Poplack, 1978, 1980). Code-switches may occur at phrase boundaries where the grammars of both languages overlap (see the switch between English and Spanish in sentence 1; Belazi et al., 1994). Code-switches may also occur within a single phrase (see the switch between Italian and German in sentence 2; Cantone & MacSwan, 2009). Code-switching at each of these syntactic locations is possible, but it is not yet entirely clear which location is more common in bilinguals' speech.

(1) The student brought the homework *para la profesora*
The student brought the homework for the teacher

(2) una *Gegend* fredda
a region cold

Code-switching also happens for many different reasons, including effects of community, discourse, and interlocutor. First, code-switching practices and patterns vary across communities and often reflect group values and norms (Heller, 2010; Myers-Scotton, 2017). Thus, code-switching can be used to strengthen the sense of community identity when speakers use it to indicate membership in the same group (Myers-Scotton, 2017; Nilep, 2006). Second, bilinguals may code-switch to produce different effects of discourse. For example, bilinguals sometimes code-switch to be better understood (Heredia & Altarriba, 2001), to change topic (Blom & Gumperz, 1972; Gumperz, 1982), or to offer a direct quotation (Ritchie & Bhatia, 2012). Last, bilinguals may decide to code-switch based on who they are conversing with. Code-switching is more common with known interlocutors, such as friends and family, than unknown interlocutors (Dewaele, 2010; Dewaele & Li, 2014). While code-switching happens for a variety of reasons, many bilinguals often cannot explicitly state why they produced a particular code-switch (Gumperz, 1982). Thus, researchers must often rely on recordings or transcripts and contextual cues to draw conclusions about the motivations behind code-switching.

The research on the production of code-switching to date has overwhelmingly focused on code-switching in conversations between two or more adults, but this is not the only time code-switching happens in bilingual settings. Bilingual parents also have the ability and opportunity to code-switch when speaking to their child(ren). This type of code-switching is relatively understudied, but some research has begun to document where parents code-switch syntactically and why they code-switch. Parents have been found to switch both intersententially and intrasententially when playing with their child in the laboratory, with intersentential code-switches occurring more frequently (Bail et al., 2015). In this study, the exact syntactic location of the intrasentential code-switches was not examined in fine-grained detail, so it is unclear if parents code-switch more between syntactic phrases or within a syntactic phrase.

The reasons that parents code-switch has received somewhat more attention. Parents have been found to code-switch for reasons such as attracting their child's attention, increasing understanding, and teaching vocabulary (Bail et al., 2015; Byers-Heinlein, 2013; Goodz, 1989). However, the frequency that parents code-switch for each of these reasons has not yet been investigated. Given that parents' speech impacts their child's language development, a thorough understanding of parents' code-switching is needed to better understand bilingual language development. Thus, Chapter 3 investigates parental code-switching, including its frequency, syntactic location, and apparent reason.

1.3.2 Code-Switching Comprehension

When bilinguals hear code-switching produced in their environment, they need to be able to understand it. Language comprehension is already a complex process when only a single language is involved. Code-switching further complicates this process by adding an additional language. How bilinguals navigate this unique task is receiving increased attention. However, again, we see that the majority of the literature about code-switching comprehension has focused on adults.

Code-switching has generally been found to be more difficult for listeners to process than single-language speech. Some have proposed that code-switching leads to processing difficulties, because the listener must switch between their two linguistic systems (Green, 1998). Following this account of bilingual language comprehension, bilinguals inhibit the language that they are not hearing (Kroll & Dussias, 2004; Macizo et al., 2010; Shook & Marian, 2013). When they encounter a switch between their languages, they must then activate the language they were previously inhibiting. The time it takes for this reactivation to occur is believed to be the underlying source of delays in the comprehension of code-switching, compared to single language processing. While it was previously believed that all code-switching was difficult for bilinguals to process, recent work has shown that not all code-switches are equally difficult and that not all bilinguals process code-switching in the same way (Gosselin & Sabourin, 2021; Ng et al., 2014; Valdés Kroff et al., 2018; Zeller, 2020).

One factor that appears to influence processing and comprehension is syntactic location. Many studies focus on code-switches that happen at a noun (e.g., "Find the *chien* [fr. dog]!; children: Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019; adults: Fernandez et al., 2019; Litcofsky & Van Hell, 2017; Tomić & Kroff, 2021), yet, as discussed above, this is not the only syntactic location that code-switching can occur. When listening to code-switches that occur intersententially (e.g., "That one looks fun! *Le chien* [fr. the dog]!"), neither French–English bilingual toddlers nor adults experience any processing difficulties or reduced comprehension compared to single language utterances (e.g., "That one looks fun! The dog!"; Byers-Heinlein et al., 2017). Similarly, Spanish-English bilinguals more readily process and understand code-switches that occur at a frequent syntactic location (e.g., before a verb phrase, "*los senadores* [sp. the senators] <u>have requested</u> the funds") than code-switches that occur at an infrequent syntactic location (e.g., within a verb phrase, "*los senadores* <u>han</u> [sp. the senators <u>have</u>] <u>requested</u> the funds"; Valdés Kroff et al., 2018). These findings suggest that the impact of syntactic location on comprehension is complex, and that further investigation is needed to better understand this relationship.

While most work investigating the comprehension of code-switching to date has focused on the impact of syntax, more recent research has begun to investigate how the content and/or function of the code-switch may influence comprehension. One recent study found that German–Russian bilingual adults process code-switches that occur at open class words (e.g., nouns) compared to closed class words (e.g., prepositions) differently (Zeller, 2020). Another study found that Spanish–English bilingual adults process two different types of open class words, nouns and verbs, differently (Ng et al., 2014). The authors of both studies argue that the different functional role these different words play impacts how they are processed.

Not only do features of the code-switch itself influence processing and comprehension, but listeners with different language experiences may process the same code-switch differently.

For example, it has been proposed that the amount of code-switching a bilingual hears impacts their ability to process code-switching efficiently. Bilinguals who are more exposed to code-switching, such as Spanish–English bilinguals in the United States, have been found to process code-switches more easily than bilinguals who are not exposed to high rates of code-switching, such as Spanish–English bilinguals in Spain (Valdés Kroff et al., 2018). Even when looking within the same population, French–English bilinguals in Canada who reported code-switching frequently in their own speech did not experience processing costs whereas those who reported code-switching infrequently did (Gosselin & Sabourin, 2021). Another individual factor that has been found to influence code-switching comprehension is language dominance. Spanish–English bilingual toddlers process code-switches from their non-dominant to their dominant language more easily than code-switches from their dominant to their non-dominant language (Bultena et al., 2015; Potter et al., 2019). These findings highlight that not all bilinguals are the same and that nuances in bilinguals' experiences should be integrated into research on code-switching comprehension.

Chapter 4 investigates how different factors, such as syntactic location and individual differences, impact bilingual children's comprehension of code-switching at a novel syntactic location, a prenominal determiner-adjective pair (e.g., Can you find *le bon* [fr. the good] duck?).

1.4 Dissertation Research Objectives

Bilingual experiences are diverse. Accurately describing and capturing this variation is crucial to fully understanding bilingual language development and comprehension. Thus, the goal of this dissertation is twofold.

The first goal is to evaluate current approaches to defining and modeling bilingualism in the field. Chapter 2 reviews current practices for operationalizing bilingualism and proposes two novel approaches drawn from the field of psychometrics. These models can help reconcile the widely used categorical approach and the recent proposals for a continuous approach. This chapter discusses how to create each model, how to analyze data with the output from each model, and how to use a model created by another researcher for a new study. Recommendations for pre-registering model creation or selection and transparent reporting practices are also provided.

The second goal is to examine the code-switching present in young bilinguals' environment and how code-switching impacts their language comprehension. Chapter 3 addresses the first component and examines the code-switching that bilinguals hear in their daily life. Using daylong, naturalistic audio recordings, parental code-switching in the home is analyzed by quantifying the frequency, syntactic location, and apparent reasons motivating parental code-switching. Chapter 4 builds on the results of Chapter 3 and addresses the question of how young bilinguals process and comprehend code-switching. Specifically, bilinguals' comprehension of code-switching at a previously unstudied syntactic location, a prenominal determiner-adjective pair (e.g., "Can you find *le bon* [fr. the good] duck?"), is investigated. In the Discussion section (Chapter 5), the relationship between code-switching input and comprehension is discussed in light of language separation and overall language development.

Combined, this dissertation contributes to the knowledge of bilingualism, particularly through a developmental lens. By focusing on code-switching, this dissertation advances understanding on how a feature unique to bilinguals' experience impacts children's language development.

2 Why Not Both? Rethinking Categorical and Continuous Approaches to Bilingualism

2.1 Introduction

Bilingualism is a complex construct that has been redefined over the past several decades. Scholars once defined bilinguals exclusively as a small group of speakers who were perfectly "balanced" in both of their languages (Lambert et al., 1959). The definition of bilingualism has since expanded to include speakers with varying degrees of proficiency and different language experiences. This change is reflected by more than 100 different group labels for bilinguals identified in the literature, such as "fully bilingual," "English Language Learners," and "successive bilingual Turkish-speaking children" (Surrain & Luk, 2017). As the definition of bilingualism evolves, models of bilingualism and the corresponding statistical techniques must develop as well. Traditionally, researchers have used a categorical approach to conceptualize bilingualism with analyses focused on the comparison of discrete groups of individuals (e.g., monolinguals and bilinguals). However, recent proposals in the literature suggest that instead of creating discrete groups, bilingualism should be modeled and analyzed as a continuous construct (e.g., Baum & Titone, 2014; de Bruin, 2019; Luk & Bialystok, 2013). This proposal has important consequences for how bilingualism is conceptualized in theory and how data are analyzed, but should bilingualism researchers abandon a categorical approach entirely? Are there ways for bilingualism to be defined and modeled beyond strictly categorical or continuous approaches? Drawing from recent advances in psychometrics and latent variable models, this chapter introduces models that integrate both categorical and continuous properties and then discusses how researchers can use these models to address complex questions in the field of bilingualism.

2.2 Current Models and Definitions of Bilingualism

An individual's bilingual status is not a trait that can be directly measured: bilingualism cannot be determined in the same way as someone's height, for example. In the psychometrics literature, a construct like bilingualism that can only be measured indirectly and is theoretical in nature is referred to as a latent construct. When measuring bilingualism, researchers often rely on a combination of observable indicators, such as language proficiency and exposure to determine an individual's bilingual status (Anderson et al., 2018; Li et al., 2006, 2014; Marian et al., 2007; Marian & Hayakawa, 2021). The use of multiple measures when evaluating an individual's bilingual status indicates that researchers (at least implicitly) view bilingualism as a multidimensional construct, or a construct comprised of "a number of interrelated attributes or dimensions" (Law et al., 1998, pg. 741). Given that the construct of bilingualism is both latent and multidimensional, deciding how to combine multiple, observable measures into one parsimonious model is a crucial step in theory development and data analysis. Multidimensional constructs most frequently follow either categorical or continuous models, depending on the theoretical relation between a latent construct and its observable measures (Diamantopoulos et al., 2008; Law et al., 1998; Meehl, 1995; Polites et al., 2012; Waller & Meehl, 1998). In the field of bilingualism, researchers frequently use a categorical model, but more are turning to continuous approaches based on recent theoretical perspectives.

2.2.1 Categorical Model

Much of the early literature on bilingualism followed a categorical model and compared bilinguals and monolinguals as discrete groups (see Figure 2.1). For example, a seminal study by Peal and Lambert (1962) compared "balanced" bilingual and monolingual children on several measures of intelligence and achievement, and the results dispelled the myth that bilingualism was detrimental to children's development. In another classic study, Ianco-Worrall (1972) found that bilingual children, defined as those who were exposed to two languages regularly and who demonstrated competence in those languages, realize the arbitrary nature of the mapping from a word's sound to its meaning earlier than monolinguals, suggesting bilinguals have advanced semantic knowledge. The comparison of bilinguals and monolinguals has also been used in more contemporary research, and a large number of studies have found differences in group comparisons of monolinguals and bilinguals, across cognitive (Bialystok, 2004; Costa et al., 2009; Prior & Macwhinney, 2010; Zirnstein et al., 2018), neuroscientific (see Del Maschio & Abutalebi, 2019; Pliatsikas & Schweiter, 2019 for reviews), and linguistic domains (e.g., Byers-Heinlein et al., 2010; Kaushanskaya & Marian, 2009; Sebastián-Gallés et al., 2012), amongst many other subfields of study.

When these bilingual and monolingual groups are examined more closely, however, variation within each group becomes apparent. For instance, bilinguals may have different ages of acquisition, language combinations, and/or degrees of proficiency, and monolinguals may have different amounts of exposure to a second language across their lifespan (e.g., many researchers consider adults to be monolingual even if they had some foreign language education in school). Researchers have recognized that the heterogeneity within the traditionally-defined bilingual and monolingual groups could obscure differences in performance within each of these groups (e.g., Abutalebi & Rietbergen, 2014; Baum & Titone, 2014; de Bruin, 2019; DeLuca et al., 2019; Luk, 2015; MacCallum et al., 2002). In order to accommodate the variation within groups and gain a deeper understanding of bilingualism, many researchers use more nuanced bilingual groups, such as "early bilinguals," "French–English bilinguals," and "nearly balanced bilinguals" (see Figure 2.2; Surrain & Luk, 2017). With the increased number of bilingual groups, researchers can compare different groups of bilinguals to each other. This allows a categorical model of bilingualism to be used to address a wide variety of research questions across subfields of bilingualism research, from infancy (e.g., Bosch & Sebastián-Gallés, 1997b) to older adulthood (e.g., Bialystok, 2004), addressing questions ranging from language development (e.g., Müller & Hulk, 2001) to cognitive benefits (e.g., Costa et al., 2009). This practice has allowed for a wide variety of comparisons to be made between bilinguals and monolinguals, as well as between different types of bilinguals, and has generated a large amount of knowledge on bilingualism.

While increasing the number of bilingual categories better captures the variability in bilinguals' experiences and abilities, categories are often poorly defined in research articles, limiting the interpretability of results (de Bruin, 2019; Hulstijn, 2012; Lehtonen et al., 2018; Surrain & Luk, 2017). This lack of clarity can be attributed to the wide variety of measures used to categorize participants and arbitrary cutoffs that may differ from study to study. Currently, there are many ways that researchers evaluate an individual's bilingual status. For example, there are several different questionnaires available to assess an individual's language background, some of which were designed for use with adult samples (LSBQ, Anderson et al., 2018; LHQ, Li et al., 2006, 2014; LEAP-Q, Marian et al., 2007), while others were designed for use with infant and/or child samples (LEQ, Bosch & Sebastián-Gallés, 2001; MAPLE, Byers-Heinlein et al., 2020; LEAT, DeAnda et al., 2016; ALDeQ, Paradis et al., 2010; BiLEC, Unsworth, 2013). While these questionnaires have similar measures, they are not identical. It would therefore be hypothetically possible that an individual could be placed into a different language group based on which questionnaire is used. Even if the same questionnaire is used across studies, the information gathered may not be used in the same way if each study prioritizes different components of a questionnaire (e.g., focusing on age of acquisition vs. frequency of use in the home).

Additionally, groups are often formed based on different cutoffs, often due to the nature of the sample available, which have varying levels of empirical support. For example, a single study may compare a group of early- and late-bilinguals, but the definition of who qualifies as an early- versus a late-bilingual may vary across studies. To illustrate, Tao and colleagues (2011) placed bilinguals into the early or late group if their second language exposure began before the age of 6 years or after the age of 12 years (respectively), whereas Baker and Trofimovich (2005) placed bilinguals into the early or late group if their second language exposure began before the age of 13 years or after the age of 15 years (respectively). Therefore, even if studies use the same labels for their bilingual groups, the groups may have different characteristics, making it difficult to synthesize findings. Because researchers cannot rely on the particular labels used in one study when comparing across multiple studies, extensive details on the bilingual sample(s) in a given study are necessary for results to be interpreted within the context of the literature.

In addition to being difficult to synthesize across studies, categorizing participants into discrete groups could have unintended consequences for statistical analyses and replicability. First, conducting group analyses when the variable of interest is actually continuous reduces statistical power and increases the chance of a Type I error (Altman & Royston, 2006; Cohen, 1983). Second, categorization could limit the reproducibility of the results if groups are formed based on an individual sample (e.g., median split), as the groups would then be quantitatively different across studies (Altman & Royston, 2006). Lastly, if groups are formed based on values of a continuous measure, a large amount of information and variability from that measure can be lost when such groups are formed (MacCallum et al., 2002). For example, if a sample of bilinguals is split based on participants' age of acquisition, there will be "early" and "late" learners. This reduces the variability within age of acquisition, and the individual ages for each participant are effectively lost. Moreover, if the split is made at an arbitrary cutoff point (say the median age of acquisition of 10 years), then those with an age of acquisition of 9 and 11 years are placed in different groups even though they may be more similar to each other than to other members of their group (i.e., an age of acquisition of 9 years is more similar to that of 11 years than that of 1 year; Altman & Royston, 2006; MacCallum et al., 2002). In sum, dividing bilinguals into groups when the underlying construct is continuous has statistical consequences and could obscure our understanding of bilingualism.

2.2.2 Continuous Model

In order to account for the full spectrum of bilinguals' experiences and abilities, some scholars have proposed that bilingualism should be viewed and analyzed as a continuous variable (Baum & Titone, 2014; de Bruin, 2019; Kaushanskaya & Prior, 2015; Marian & Hayakawa, 2021; Takahesu Tabori et al., 2018). Under such an approach, the continuum would span the range from completely monolingual (i.e., never having any exposure to a second language) to fully proficient bilingual (i.e., "balanced;" see Figure 2.3). It would be possible to create a continuum of bilingualism based on a single variable (e.g., years spent speaking two languages). However, given that bilingualism is a latent and multidimensional construct, using a variety of measures might better place individuals on a bilingualism continuum. These different measures will need to be mathematically combined into a final bilingualism score. For example, the concept of language entropy incorporates participants' responses to questions about their language exposure, language proficiency, language use in different contexts, and L2 accent perception on a single continuous scale (Gullifer & Titone, 2019). When using a continuous approach, scholars will need to determine which measures to include and how they will be algebraically combined to result in a final bilingualism score (Law et al., 1998), for example giving more weight to some dimensions (e.g. age of acquisition) than others (e.g., time spent listening to the radio in the second language). Marian and Hayakawa (2021) have recently dubbed this type of standardized bilingualism index a "Bilingualism Quotient." It is important to note that the relationship between different measures and the final bilingualism score does not need to be linear. For instance, age of acquisition could follow a pattern of non-linear decrease resembling threshold effects seen in sensitive periods for language development (Werker & Hensch, 2015; Werker & Tees, 2005).

A continuous model would allow researchers to investigate subtle effects of bilingualism and would therefore be useful in specialized applications. For instance, the investigation of potential cognitive benefits of bilingualism in adults could benefit from the ability to detect smaller effects, and using a continuous model could potentially establish thresholds to see effects of bilingualism in this domain (e.g., Cummins, 1976; De Cat et al., 2018; Ricciardelli, 1992).

While using a continuous model for bilingualism may be appropriate in some research domains, it is unlikely that this model will become the standard across all subfields of bilingual research, as the benefits may not apply to certain subfield-specific contexts. For example, some subfields (e.g., research with special populations such as infants, or children with developmental delays) will tend to focus on large effects in smaller samples, making a continuous model less practical than a categorical approach. Moreover, categorical approaches might be more appropriate than continuous ones in some research contexts, for example enrolment in a language immersion program is inherently categorical (i.e., children are or are not enrolled), a point that we will return to later in this chapter. Therefore, both continuous and categorical models may be useful in advancing bilingualism research depending on the particular study.

2.3 Expanding Models of Bilingualism

Both categorical and continuous models of bilingualism have their advantages and disadvantages. Categorical models are easy to interpret, but the groups used in the analyses may be heterogeneous. Continuous models accommodate more individual variation but may not be practical in all bilingualism research and may be inappropriate if the underlying construct is actually discontinuous. Each one can answer different research questions, but given that bilingualism is a complex construct, some research questions may be best addressed by some combination of the two. Are models available that better reflect the complexity of bilingualism by incorporating the advantages of both categorical and continuous models? Other areas of research, such as psychometrics, may offer innovative solutions to defining and modeling bilingualism (Borsboom et al., 2016). While there are many different psychometric models that bilingualism researchers can consider, here we introduce two interesting possibilities: the factor mixture model and the grade-of-membership model. Like current approaches to modeling bilingualism that rely on participants' responses to a series of questionnaires or tasks, both of these models find patterns within participants' responses about their language history, proficiency, and any other variables relevant to defining bilingualism (Andreotti et al., 2009; Clark et al., 2013; Masyn et al., 2010). Additionally, researchers can decide which participant data is of theoretical interest to include in the model (e.g., language attitudes, proficiency, age of exposure). Unlike current approaches, categories are not pre-defined by the researcher, nor are they formed by potentially arbitrary cutoffs determined by the researcher. Instead, categories emerge as clusters based on statistical patterns in the data. Furthermore, each of these models offers the possibility of analyzing data continuously, which could increase statistical power of analyses involving the dependent variable if bilingualism does exist on a continuum (Altman & Royston, 2006; Cohen, 1983). In sum, each of these models is more comprehensive than current research practices and would allow researchers to incorporate both categorical and continuous properties when analyzing their data.

2.3.1 Factor Mixture Model

Factor mixture models are based on the idea that variation can exist within categories (Lubke & Muthén, 2005; McLachlan & Peel, 2004), thus individuals are both placed into separate categories and given a score on a continuous scale (Clark et al., 2013). Depending on the constraints set when developing the model, this continuous score could be interpretable relative to all participants, or only relative to participants within the same category. For an example unrelated to bilingualism, children could be divided into categories based on whether or not they have a conduct disorder, and the degree to which they exhibit symptoms is allowed to vary within each group (i.e., children in the group with conduct disorders vary in severity of symptoms; Clark et al., 2013).

With the definition of bilingualism expanding beyond the view that only individuals who are "balanced" in both of their languages are bilingual, there is inherently more variation across individuals who would now be considered bilingual. Factor mixture models could capture the variation within bilinguals by classifying participants into either a monolingual or bilingual group and accounting for variation within each of those groups (see Figure 2.4). Factor mixture models can also accommodate multiple groups. Allowing multiple bilingual groups in a factor mixture model could potentially mirror groups that already exist in the literature (e.g., simultaneous, sequential, etc.), and subsequently capture the heterogeneity within those groups (Clark et al., 2013; Sulpizio et al., 2020). While theory can drive the number of categories and the measures that are included in a final bilingualism score, it should be noted that the number of groups and the way that different variables contribute to the continuous score are typically determined through an iterative modeling process. In this process, the number of groups and how different variables define group membership are systematically varied to find the strongest factor mixture model, although the researcher can set theoretically-motivated constraints on models that will be considered (Clark et al., 2013; Nylund et al., 2007).

For a concrete example, imagine Dr. Factor-Mixture who is working on a project investigating the potential effect of bilingualism on a memory task and plans to use a factor mixture model to identify bilinguals and monolinguals in her research. Dr. Factor-Mixture collects information from 150 participants – the minimum recommended sample size for creating a factor mixture model (Lubke & Neale, 2006) – about their language experience and history via the Language History Questionnaire (LHQ; Li et al., 2014) before they complete the memory task. Once all her data are collected, she uses the participants' responses to the questionnaire to determine their bilingual status. She will use the FactMixtAnalysis package (Viroli, 2012) in R, her preferred statistical software (although she could have also used Mplus; Muthén & Muthén, 2016). Using the observed patterns of responses to the questionnaire, participants are placed into different groups and within each group are given a composite, final score on a continuous scale indicating how they are situated within the group (Clark et al., 2013; DiStefano et al., 2009). Dr. Factor-Mixture can choose a specific type of factor mixture model that either uses the same or different variables to determine continuous scores in each group depending on her research goals and theoretical conceptualization of bilingualism (Clark et al., 2013). Dr. Factor-Mixture expects that there may be different types of bilinguals in her sample (i.e., sequential and simultaneous bilinguals), so she runs models with different numbers of expected groups. In order to compare the goodness of fit for different models and identify the most parsimonious model, Dr. Factor-Mixture compares the AIC and BIC values of each model and selects the one with the lowest value (Hallquist & Wright, 2014). These values indicate how closely the data fit a particular model. When comparing the results, the model that contains 4 groups built from different variables for each group is the most parsimonious and is selected as the final model. When Dr. Factor-Mixture examines the output of the final model, she looks at how different variables contribute to group membership and sees that these groups could be described as monolingual, sequential low-proficiency bilingual, sequential high-proficiency bilingual, and simultaneous high-proficiency bilingual. Dr. Factor-Mixture can now analyze the participants' scores from the memory task categorically using the groups identified in the model in an ANOVA or use a regression model to additionally incorporate participants' continuous scores within each group.

2.3.2 Grade-of-Membership Model

Grade-of-membership models also allow for variation within categories. Such models place individuals into different categories, but uniquely allow for individuals to simultaneously belong to different categories to varying degrees (Andreotti et al., 2009; Erosheva, 2005). Some individuals overwhelmingly belong to one group, and the model consequently places them into that group. Some individuals may be somewhere in between multiple groups, belonging to different groups to different degrees. Grade-of-membership models capture in-between cases, where individuals' categorization is not as clear, through a "fuzzy set." This set has no definitive boundaries, and individuals belong to this set to different degrees. Grade-of-membership models can accommodate multiple groups and the overlap between them. For an example unrelated to bilingualism, individuals can be simultaneously affiliated with different political parties, because their ideologies fall somewhere in between those most characteristic of the different groups (Gormley & Murphy, 2009).

When applied to bilingualism, a grade-of-membership model could still include monolingual and bilingual groups but would also accommodate individuals who do not necessarily fit strict definitions for either group (see Figure 2.5). Imagine an individual who studied a second language for several years and obtained an intermediate level of proficiency, but who no longer uses the language frequently. They might not qualify as either monolingual or bilingual by the definitions used in many studies. Individuals like this have often been less studied in the literature. However, it might still be important to include these individuals in studies in order to gain a more comprehensive view of bilingualism. Therefore, incorporating a grade-of-membership model and the "fuzzy set" between different groups of bilinguals and monolinguals could offer more insight into how language experience influences a wide variety of factors.

To see this in practice, imagine Dr. Grade-O'Membership who is investigating the effect of bilingualism on word learning in adults. Dr. Grade-O'Membership recruited 200 participants – the minimum recommended sample size to allow for accurate group identification in grade-of-membership models (Holmes Finch, 2020) – and asked his participants extensive questions about their language history and proficiency using the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007). He decides to analyze the responses to these questions using Mplus (Asparouhov & Muthen, 2006; Muthén & Muthén, 2016), but he could have also used the mixedMem package in R (Wang & Erosheva, 2015). Dr. Grade-O'Membership builds several models with different number of groups and selects the final model, which happens to have only 2 groups, by identifying the model with the lowest truncated sum of squared Pearson residuals (X^2_{tr}; Erosheva et al., 2007; Holmes Finch, 2020). Based on their responses, each participant is given a probability of belonging to each of the 2 groups identified in the sample; the total of these probabilities will sum to one. Dr. Grade-O'Membership can determine if a participant should be placed in the bilingual or monolingual group, based on the group the model says they have the highest probability of belonging to. He notices that very few participants have intermediate probabilities, so decides that his sample has more of a categorical structure. He then determines which group learned more words using a 2-sample *t*-test. He could also use the probability that each participant belongs to the bilingual group to analyze the data continuously and examine the relationship between the degree of bilingualism and the number of words learned using a regression model.

2.3.3 Incorporation of New Models

Both the factor mixture model and the grade-of-membership model are tools that researchers can use to better represent the underlying structure of bilingualism and better address questions in the field. They could be incorporated into research on bilingualism by following several steps. In order to benefit from either of these comprehensive model approaches, a new model will first need to be created and validated following the steps explained in the hypothetical examples above. This would involve creating new datasets or using pre-existing databases with information about a wide range of bilinguals and monolinguals on a variety of bilingualism measures, such as language proficiency and history (e.g., via an extensive questionnaire such as the LHQ, Li et al., 2014; or LEAP-Q, Marian et al., 2007). Then various iterations of either the factor mixture or grade-of-membership model would be built and evaluated for goodness of fit using statistical software (Clark et al., 2013). Once a parsimonious model has been fit to the data, researchers can use the model to address a variety of research questions. Researchers can use models that they have built themselves or models built by other researchers. If several studies addressing the same question use the same model, researchers will be able to make direct comparisons across these studies.

For an example of how researchers could use previous models, take Dr. Resourceful who is studying attention. Dr. Resourceful is only able to test 75 participants, which is not an adequate sample size to develop their own factor-mixture or grade-of-membership model. Instead, they opt to use the model developed by Dr. Factor-Mixture to evaluate the bilingual status of the participants they do have, because they are studying a similar population. Dr.

Resourceful will need to give their participants the Language History Questionnaire (Li et al., 2014), so that participants answer the same questions that Dr. Factor-Mixture used to create the model, and feed participants' responses to specific items into the model. This will output a bilingualism score for each participant, as well as identifying which of the 4 groups from the original model the participant belongs to. Dr. Resourceful discovers that none of their participants are placed into the sequential high-proficiency group but are split relatively equally into the remaining groups. Because each of the groups has different variables contributing the bilingualism score (due to the nature of the original model developed by Dr. Factor-Mixture), a continuous analysis of all participants is not possible in this model, but Dr. Resourceful can approach their analyses in one of two ways. They can analyze the data through a categorical lens, using the monolingual, sequential low-proficiency, and simultaneous high-proficiency groups formed by the model, or they can incorporate both the categorical and continuous information from the model in the analyses by computing a separate regression model using the final bilingualism score for each of their groups.

The factor mixture and grade-of-membership models are simply two of many models that researchers could consider employing in the field of bilingualism. If we look to the field of psychometrics, there are a wide variety of models that could help researchers better define and model bilingualism, such as different forms of factor analysis (Anderson et al., 2018) or cluster analysis (Woodbury & Manton, 1989). In using more complex models, information on modeling decisions will need to be made explicit, and assumptions about the nature of bilingualism could ultimately be challenged. By addressing these issues in the field, researchers will be able to drive theories of bilingualism forward. While these complex models will help to operationalize bilingualism, it is necessary to address how to best incorporate them into the field.

2.3.4 Standardization in the Field

When moving towards more comprehensive models of bilingualism, some may argue that there is a single best model of bilingualism that should be used in the field, including across different subfields and studies (Marian & Hayakawa, 2021). However, this approach could face obstacles in the measures that are available across the stages of development and the statistical analyses that can be conducted with different populations. Additionally, standardization within the field of bilingualism could limit the number and type of research questions that can be addressed.

First, a standard definition of bilingualism may be difficult to implement across different populations and stages of development. For example, it is possible to gather a wide range of data on an adult's language proficiency and background through questionnaires or language tests (Anderson et al., 2018; Li et al., 2014; Marian et al., 2007; McNamara, 2000). This provides a comprehensive view of an individual's language experience that could be used in analyses. However, gathering the same in-depth information on an infant's language experience is much more difficult. Infants are unable to respond to direct questions, so their caregivers must provide information about their language experience, which is often limited to information about their language exposure (Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., 2020). Trying to use the same standardized measure for both adults and infants would be ineffective and ultimately unsuccessful. We argue instead that in order to increase transparency, bolster comparisons across studies, and help replication efforts, researchers should include detailed descriptions of their definition, measures, and model of bilingualism (Esposito et al., 2015; Luk et al., 2017). Furthermore, where possible, researchers who work with similar populations should try to reach a consensus on using a single measure (De Cat et al., 2021).

Second, bilingualism may have a different underlying structure in different target populations or in the context of different research questions, and, as discussed above, it is important that statistical analyses accurately reflect this underlying structure (Altman & Royston, 2006; Cohen, 1983; MacCallum et al., 2002). For example, in a study investigating if there is a difference in bilinguals' and monolinguals' ability to discriminate two languages in infancy (Byers-Heinlein et al., 2010; Nazzi et al., 2000), a categorical construct such as language group (i.e., monolingual vs. bilingual) might appropriately characterize the sample, and *t*-tests, ANOVAs, or regressions with categorical predictors would be appropriate analytic approaches. By contrast in a study investigating how bilingual experiences (e.g., age of acquisition of their second language) affect brain function (DeLuca et al., 2019), participants might be best characterized in terms of a continuous measure of bilingualism, and correlations or regression models would be appropriate. Finally, as this paper has proposed, in many cases the sample might have both categorical and continuous characteristics, for example in a study of undergraduate students who come from diverse monolingual and bilingual backgrounds and have different language histories. Here, either a factor mixture or grade-of-membership model could be appropriate. Because of the variety of samples and research questions in the field of bilingualism, it is important that a variety of models be accepted in the field and for researchers to carefully consider which model best addresses their population and research question.

2.3.5 Future Directions

This paper has discussed four different models of bilingualism that scholars have used or could use in their research. The traditional practice of using a categorical model and the recently proposed continuous model of bilingualism are the tip of the iceberg for how bilingualism can be defined and modeled. We have suggested two other types of models for bilingualism researchers to consider: the factor mixture model and the grade-of-membership model. These models extend the current thinking about how bilingualism should be defined and understood, as they incorporate both categorical and continuous aspects.

Although the aim of this paper is to encourage researchers to consider different models of bilingualism, we caution against too many models being used across the literature. We recommend that particular subfields compare the relative theoretical and practical merits and performance of different models, and carefully consider the types of participant data used to create their models (e.g., questions about language proficiency and use versus questions about language attitudes). Ideally, subfields will converge on the model that is most appropriate for their research questions and populations and converge on a standard approach to collect such data (e.g., a consistent questionnaire). For the researchers who are developing models, we encourage them to pre-register the steps that they will take and the comparisons that they will make to arrive at the final model, including the number of different groups and the combinations of variables they will try. Once the model has been finalized, researchers can transparently report the creation and selection process and share their scripts, so others can use the same model. Similarly, for researchers who are using previously developed models, we suggest that they consider which model to use based on their research question and the typical models used in their subfield before data analysis begins and to pre-register this choice, as well as their commitment to use the same materials that were used in the development of the model. This will reduce the chances of *p*-hacking and tinkering with group definitions until results are statistically significant or match the original hypothesis, which can increase Type I error and lead to less robust results (Simmons et al., 2011). We also encourage all researchers to share their data to increase transparency and contribute to standardization efforts. Combined, taking these steps will help a particular subfield converge upon a single model best suited to its needs. Adopting more nuanced models will ultimately allow for a wider range of research questions to be addressed and for advancement of theories of bilingualism.

Figure 2.1: Representation of a categorical model of bilingualism.

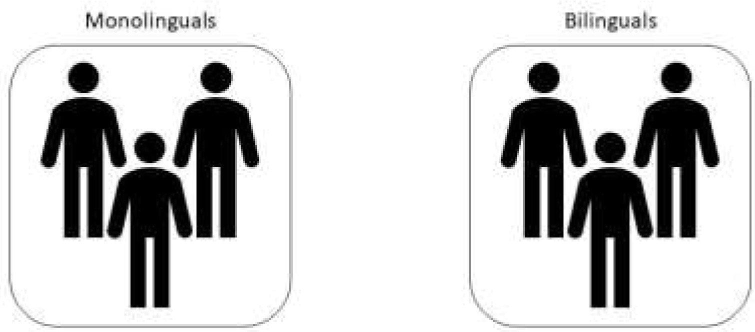


Figure 2.2: Representation of a categorical model of bilingualism with many different possible groups of bilinguals.

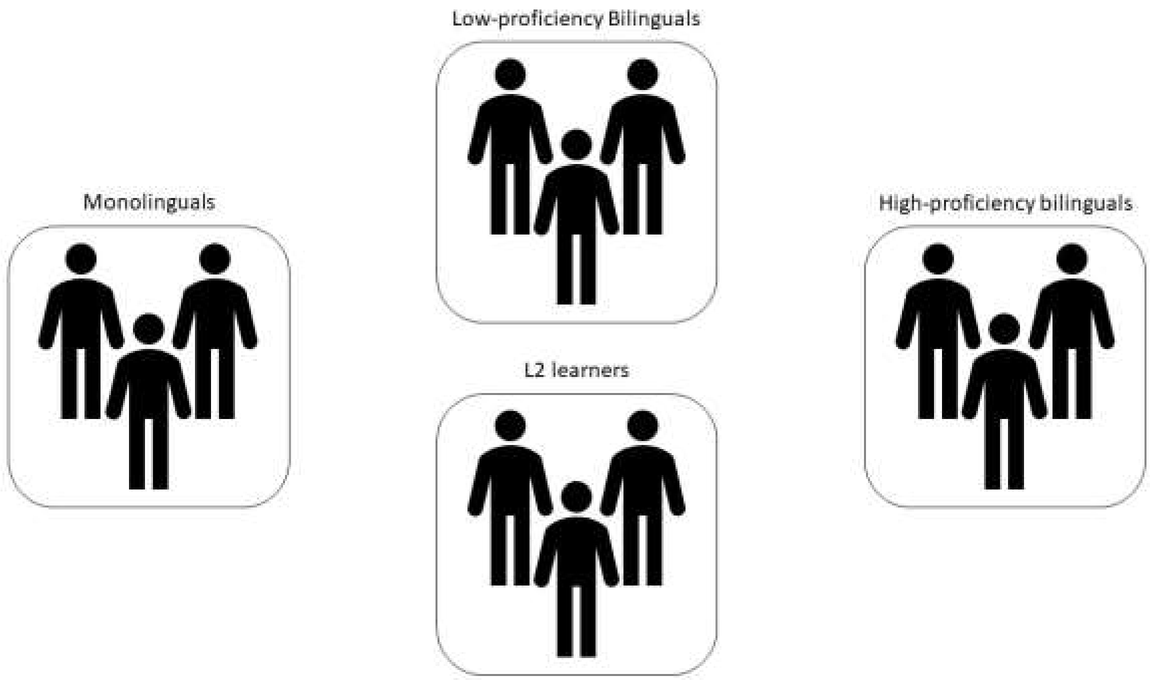


Figure 2.3: Representation of a continuous model of bilingualism.

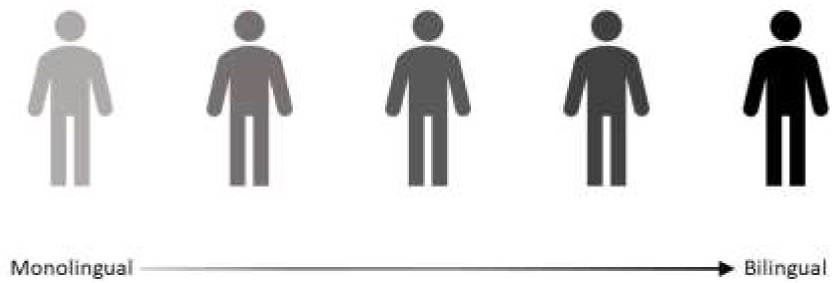


Figure 2.4: Representation of a factor mixture model of bilingualism where data can be analyzed based on categorical membership or placement on a continuum.

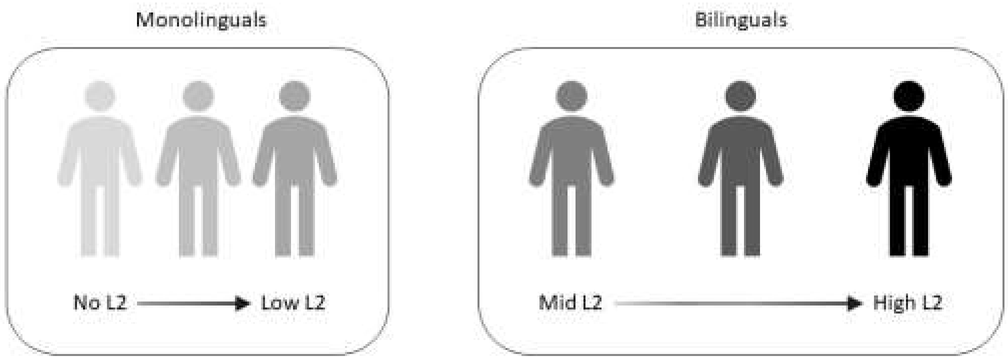
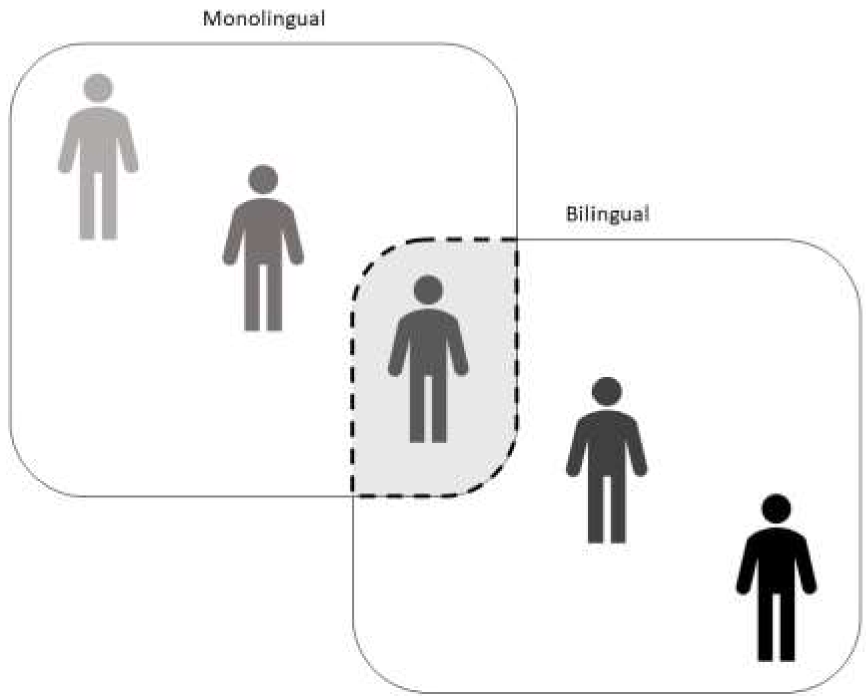


Figure 2.5: Representation of a grade-of-membership model of bilingualism where data can be analyzed based on categorical membership or placement on a continuum.



3 Code-Switching in Parents' Everyday Speech to Bilingual Infants

3.1 Introduction

In environments where multiple languages are used, bi- and multilingual speakers can combine more than one language in their conversations, a phenomenon known as code-switching. Code-switching was originally believed to be the result of a language deficiency (Weinreich, 2010), a strategy used by bilinguals to compensate for a lack of proficiency in either one or both languages (Heredia & Altarriba, 2001). Others believed that code-switching threatened the "purity" of a language (Myers-Scotton, 2017). Scholars today, however, overwhelmingly reject these views and recognize that code-switching is a systematic and complex linguistic phenomenon that is typical of bilingual communities (Hoff & Core, 2015; Ritchie & Bhatia, 2012; Yow et al., 2018). For the past several decades, researchers have investigated how often, where syntactically, and why code-switching occurs. While this body of research is vast, it has typically focused on speech between bilingual adults. What remains largely unstudied is the nature and purpose of code-switching when bilingual adults speak to young children. Here, we analyzed the speech input of 21 French–English bilingual families in Montreal via day-long, at-home audio recordings that were made when infants were both 10 and 18 months old, a critical period for language development. Our goal was to understand the nature of code-switching in parental speech to bilingual infants, focusing on a) frequency, b) syntactic location, and c) apparent reasons for code-switching.

3.1.1 Why Caregivers' Code-Switching Matters for Understanding Language Development

Adults routinely modify their speech when interacting with children (Fernald, 1989). For example, many language communities around the world use infant-directed speech, which has characteristics that include variability in pitch (Stern et al., 1983), higher pitch (Albin & Echols, 1996), shorter utterances (Soderstrom et al., 2008), more repetition (Hills, 2013), and lengthening of final syllables (Albin & Echols, 1996). The exact qualities of infant-directed speech vary between parents and have been linked to variations in infants' linguistic abilities (see Soderstrom, 2007 for a review).

For bilingual caregivers, infant-directed speech may contain code-switches, which like other aspects of infant-directed speech, could impact language development either positively or negatively. On one hand, laboratory studies have suggested that code-switched speech can be more challenging for bilingual children to process than single-language speech (Byers-Heinlein, 2013; Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2018). Difficulties in language processing could ultimately lead to delayed language development. On the other hand, code-switching may be a useful strategy for bilingual caregivers to support their child's development in both of their languages. For example, long-term exposure to code-switching could prepare infants for processing dual-language input (Orena & Polka, 2019). Further, code-switching could be used to scaffold bilingual vocabulary acquisition, by providing terms in each language. However, we still have a poor understanding of how often and what types of code-switches infants encounter in their daily life. Investigating the quality and quantity of infant-directed code-switching is a crucial first step in understanding how it might affect language development.

3.1.2 Frequency of Code-Switching

Code-switching is common in bilingual and multilingual communities (Myers-Scotton, 2017), and it also occurs in parents' speech to their children (Goodz, 1989). A questionnaire-based study in Vancouver, Canada, found that more than 90% of bilingual parents (English and another language) reported engaging in code-switching when speaking to their children (Byers-Heinlein, 2013). Although code-switching was common, the frequency across parents was found to be highly individualistic. Within-sentence code-switching roughly followed a normal distribution, highlighting the variation between parents' rates of code-switching. Similarly, an observation-based study in Maryland, USA, observed that all of their Spanish–English bilingual parents used code-switching during a play session with their child (Bail et al., 2015). While code-switching occurred, on average, in 15.8% of all utterances by each parent, this ranged by parent from 0.4 to 58.5% (Bail et al., 2015). This variation across speakers is also commonly observed in studies on adults' code-switching behaviors (Dewaele & Li, 2014; Dewaele & Zeckel, 2016).

The frequency of parental code-switching may be an important factor in a bilingual child's language development. For example, children code-switch at a similar rate as their parents, suggesting that parental code-switching serves as a model (Genesee et al., 1995).

Additionally, several studies have investigated the relationship between the frequency of parents' code-switching and their child's vocabulary size, but these studies have reported divergent findings: some indicate that code-switching may negatively impact a child's vocabulary development (Byers-Heinlein, 2013) while others indicate that code-switching has no impact on a child's vocabulary development (Bail et al., 2015; Carbajal & Peperkamp, 2020).

One possible reason for these mixed results in the literature is that different studies have used different methods to measure the frequency of parental code-switching. One method that has been used is observing parents during a play session in a laboratory environment (Bail et al., 2015; De Houwer & Bornstein, 2016). This allows for direct measurement of code-switching frequency but is limited, because parents may not engage in their usual code-switching behaviors due to perceived expectations in the laboratory environment. Another method is to use questionnaires asking parents to rate the frequency of their code-switching (Byers-Heinlein, 2013; Place & Hoff, 2016). However, parents' self-reported frequency of code-switching may not reflect their actual frequency of code-switching (Bail et al., 2015). Additionally, when comparing parents' self-reported code-switching to their child's performance on language tasks, no relationship between the two measures has been observed (Place & Hoff, 2016; Schott, Mastroberardino, et al., 2021). This indicates that while self-report is time-efficient, parents may not be consciously aware of how much they code-switch, meaning this measure could be inaccurate. Lastly, rates of parental code-switching have been measured via a diary method, where parents indicate whether they spoke to their child in one or both of their languages in a given 30-minute block (Place & Hoff, 2011, 2016). This measures whether the two languages co-occur temporally across large blocks of time but does not quantify in more fine-grained detail exactly how much code-switching a child hears.

To better understand how code-switching might influence children's language development, and given that the frequency of code-switching is highly variable between individuals and difficult to measure precisely, a new approach is needed to more accurately measure the frequency of parental code-switching. One solution is to obtain recordings from parents speaking to their child in their home. This provides a more accurate picture of everyday code-switching in families; the exact number of code-switches can be counted and analyzed. This method circumvents measurement issues associated with observation in an unfamiliar laboratory environment and self-report. Moreover, it provides the opportunity to assess the accuracy of self-report measures and evaluate how these measures could be incorporated into future research. Our first research goal was to observe how frequently code-switching occurs in the daily life of bilingual families and to compare this to self-report measures of code-switching.

3.1.3 Syntax of Code-Switching

Raw measures of the frequency of code-switching do not account for the diverse syntactic locations where code-switching can occur. Code-switching is a rule-governed language phenomenon, and code-switches do not occur in random syntactic locations (MacSwan, 2012). Intersentential switches occur between sentences, and thus are not subject to syntactic constraints (e.g., Let's read a book. *Je vais lire un livre*. [fr. I'm going to read a book.]); intrasentential switches occur within a sentence and are governed by syntactic rules (e.g., I'm going to read *un livre* [fr. a book]; MacSwan, 2012). This distinction is important even in early development, as some research has suggested that 20-month-old bilinguals more readily process intersentential code-switches than intrasentential code-switches (Byers-Heinlein et al., 2017). Indeed, previous research has shown that parents tend to code-switch more between sentences than within a sentence when playing with their child (Bail et al., 2015).

Intrasentential code-switches can occur at several different locations (MacSwan, 2012). For decades, scholars have proposed various theories and rules to describe the systematic nature of code-switching (e.g., Azuma, 1992; MacSwan, 2012; Myers-Scotton, 1997; Poplack, 1980; Sankoff & Poplack, 1981; Woolford, 1983). While these theories vary on their exact rules, they generally converge on the idea that code-switching can occur when the grammars of the two languages overlap in some way (Poplack, 1978, 1980).

To our knowledge, only one study has examined the syntactic locations of intrasentential code-switches in parental speech, finding that over half of intrasentential code-switches occurred between a determiner and a noun (e.g., *el* [sp. the] apple; Bail et al., 2015). One important distinction may be whether intrasentential code-switches occur at a syntactic boundary (e.g. The student brought the homework *para la profesora* [sp. for the teacher]; example and translation from Belazi et al., 1994), or within a syntactic phrase (e.g. una *Gegend* fredda [a region cold]; Italian–German example and translation from Cantone & MacSwan, 2009). Bilingual infants show sensitivity to the syntactic structure of their languages by age 7 months (Gervain & Werker, 2013), which suggests that they might also be sensitive to the syntactic location of code-switches. Due to the potential differences in children's processing of code-switches at various syntactic boundaries, our second research goal was to investigate the prevalence of code-switches occurring at a syntactic boundary compared to ones occurring within a syntactic phrase. This is a tractable way to begin examining the effects of the syntactic properties of parental code-switching on speech processing and language development in general.

3.1.4 Reasons for Code-Switching

Speakers may engage in code-switching for different reasons depending on whether they are interacting with another adult or with a child. Certain reasons that drive code-switching in adult conversations might also apply to parent–child speech. At the same time, there may be unique motivations that parents have for code-switching that support their child's language development.

First, code-switching behaviors vary significantly between different communities of bilinguals (Heller, 2010). For many, code-switching serves to reinforce a community's identity by following the accepted local norms and functions of code-switching (Nilep, 2006). For example, the communal identity can be strengthened when a speaker code-switches in order to use an idiom from one of their languages or to share a piece of cultural wisdom or history, a type of linguistic borrowing (Ritchie & Bhatia, 2012). Parents may code-switch with their young children in the same way that they do with other community members. This could serve to teach their child their community's norms and expectations. Indeed, research on children's early productions of code-switching have found that children code-switch at a similar rate to their parents, suggesting that parental code-switching may serve as a model for developing bilinguals (Comeau et al., 2003; Yip & Matthews, 2016). Modeling norms might also occur through borrowings that are common in the community, which could include baby- and child-specific terms. For example, in Montreal, Canada, it is common for a child's stuffed toy to be referred to with the French word "*toutou*" regardless of the language that the parent is speaking. Using this term when speaking English would be an instance of code-switching.

Another important driver of code-switching in adult–adult conversations is to improve understanding of the speaker by their conversational partner (Heredia & Altarriba, 2001). Similarly, bilingual parents may use code-switching to adapt to their child's knowledge, for example code-switching to produce a word that they know their child understands rather than its unfamiliar translation equivalent. There are also reports that parents sometimes code-switch in order to teach their child a new word, again using code-switching strategically to enhance their children's comprehension and learning (Byers-Heinlein, 2013). For example, bilingual parents have been observed to code-switch in order to provide a translation from one language into the other (Bail et al., 2015). Code-switching in these circumstances may help to support children's language learning.

Finally, adults have been observed to code-switch in conversations to create metaphorical effects in the discourse (Blom & Gumperz, 1972; Gumperz, 1982; Ritchie & Bhatia, 2012), for example using direct quotations, such as, "they be like '*loca, loca*'" [sp. honey, honey] (example and translation from Bailey, 2000). Metaphorical code-switching is difficult for analysts to classify, and even native-speakers do not consciously understand all of the motivations driving metaphorical code-switching (Gumperz, 1982). Parents may also code-switch to produce metaphorical effects unique to a child's language development. This could include code-switching in order to get their child's attention, emphasize a point, or discipline their child (Bail et al., 2015; Byers-Heinlein, 2013; Goodz, 1989).

In sum, there are numerous reasons why adults code-switch in speech to other bilingual adults, as well as additional reasons why adults might code-switch when speaking to their children. The reason(s) motivating a parent's code-switching could potentially bolster a child's language development. Additionally, a single code-switch may be motivated by multiple reasons. However, there is little research that quantitatively investigates parents' motivation for code-switching when speaking to their child. Our third research goal was therefore to explore and quantify parents' apparent reasons for code-switching in speech to their young children, given the paucity of research on this topic.

3.1.5 Changes Across Development

Parents adapt their speech to their child's linguistic abilities. For example, prosodic features, such as pitch, change across an infant's first months (Kitamura & Burnham, 2003; Kitamura & Lam, 2009; Stern et al., 1983). As a second example, properties of parents' speech, such as vowel articulation (Lam & Kitamura, 2012) and syntactic complexity (Elmlinger et al., 2019), appear to change in response to their infants' feedback. This work has focused on monolingual parents, but it is likely that bilingual parents also alter their speech based on their infant's feedback. It is currently unknown whether parents' code-switching changes in response to an infant's developing language abilities, as previous studies have not investigated properties of parental code-switching longitudinally beyond whether or not code-switching occurs (De

Houwer & Bornstein, 2016). Thus, our fourth research goal was to examine how parental code-switching may change across their infant's development.

3.1.6 Current Study

The current study investigated the code-switching behaviors of parents in Montreal, Canada. Montreal is a unique environment for studying bilingualism, because both French and English are widely spoken throughout the city, and both have high status in the community. This creates a favorable environment for investigating code-switching. Below, we detail the predictions associated with each of our research questions:

RQ1a: How often do parents code-switch?

We expected to observe code-switching in all families. However, we expected that the frequency of code-switching would vary across families. Such a finding would be consistent with previous research (Bail et al., 2015; Byers-Heinlein, 2013).

RQ1b: How reliable are self-report measures of code-switching?

In addition to measuring the frequency of parents' code-switching, we had the opportunity to compare this direct observation to a self-report measure (the Language Mixing Questionnaire; Byers-Heinlein, 2013), thereby evaluating the validity of such measures.

RQ2: Where do parents code-switch syntactically?

We predicted that parents would code-switch both between and within sentences. Generally, we expected to observe more intersentential than intrasentential code-switches, as previously reported by Bail and colleagues (2015). For intrasentential code-switching, we predicted that it would more often occur between syntactic phrases than within a syntactic phrase (Woolford, 1983). This pattern may emerge because switches at a syntactic phrase boundary are easier to produce or process.

RQ3: Why do parents code-switch?

We anticipated that code-switching would occur for a variety of apparent reasons. Previous research suggests that parents may code-switch for reasons such as boosting their child's understanding, borrowing a term from the other language, providing a translation equivalent, getting their child's attention, emphasizing a point, or disciplining their child. Because previous research has not addressed how frequently parents code-switch for each of these reasons, we did not have any predictions as to which reasons would be more frequent than others or what combination of reasons may motivate a single code-switch.

RQ4: Do patterns change across the infant's development?

Due to the great advances in children's language skills between 10 and 18 months of age, we expected parents' code-switching frequency to increase between these two time points, as they adjust to their children's language skills (e.g., Stern et al., 1983). This prediction also follows from an implicit assumption that parents adapt their input to their children's language processing abilities.

3.2 Method

Data were drawn from the Montreal Bilingual Corpus (Orena et al., 2019), which contains daylong home recordings for French–English bilingual children recorded at age 10 months and again at age 18 months. We initially conducted pilot coding of data from 2 children at 10 months to verify and finalize our coding scheme. Prior to listening to or coding the remaining code-switches, we then pre-registered our methods via the Open Science Framework at https://osf.io/a52ku. Any deviations from the pre-registration are noted and justified. All data, including those from the 2 pilot children, were included in the final analysis. This research was approved by the Institutional Review Board at McGill University (IRB # A05-B20-16A).

3.2.1 Participants

Participants who contributed to the corpus were families with a young infant who heard French and English at home ($n = 21$). Infants were 10 months of age (M = 9m29d, *Range* = 9m15d – 10m14d) during their first visit to the laboratory. Most of these families returned for a second visit ($n = 16$) when infants were 18 months of age (M = 18m29d, *Range* = 18m4d – 20m26d). As reported by parents, none of the infants had an auditory or developmental neurocognitive disorder. Parents also reported being from a mid to high socioeconomic background, with a mean Hollingshead score of 52.2 (*Range* = 31 – 66 out of a possible 66).

Using a common cut off in the field of infant and child bilingualism (Byers-Heinlein, 2015), initial eligibility criteria for the corpus required that infants have at least 25% of their overall exposure to both English and French, and that they have daily exposure to both languages. Infants' language exposure was first estimated during a phone screening and then evaluated more thoroughly upon their visit to the lab with a language exposure questionnaire (LEQ; Bosch & Sebastián-Gallés, 2001) using the Multilingual Approach to Parent Language Estimates (MAPLE; Byers-Heinlein et al., 2020). Based on this questionnaire, 3 infants no longer met the language exposure criteria as they had slightly lower than the 25% minimum exposure to their non-dominant language. However, these infants were still included in the corpus as well as the current analyses because they all received daily exposure to both French and English (Orena et al., 2019). At 10 months, twelve infants were in a French-dominant environment (i.e., 56 – 79% of their language exposure was in French), and 9 were in an English-dominant environment (i.e., 55 – 76% of their language exposure was in English). Four infants also heard a small amount of a third language in the home (i.e., Arabic, Kannada, Portuguese, and Spanish), but this constituted less than 5% of each infant's language exposure. At 18 months, 8 infants were in a French-dominant environment (60 – 78% French) and 8 were in an English-dominant environment (50 – 78% English).

Each family in the corpus included two different-sex parents. While all parents reported knowledge of English and French, not all parents reported speaking both languages to their infant. Of the 42 parents in the corpus, 26 reported that they spoke both languages to their infant, while 16 reported speaking only one language to their infant. Each parent, except one, completed the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). Parents' age of acquisition ranged from 0 – 17 years old ($M = 4.78$, $SD = 4.94$) for English and from 0 – 21 years old ($M = 3.10$, $SD = 5.27$) for French. Parents also rated their proficiency for speaking, comprehension, and reading from 0 to 10 in both English and French and reported a mean proficiency score of 9.23 ($SD = 0.86$, *Range* = 6.33 – 10) in English and 9.42 ($SD = 1.12$, *Range* = 5.67 – 10) in French. All parents completed the Bilingual Dominance Scale (Dunn & Fox Tree, 2009), and reported an average dominance score of 3.48 ($SD = 12.90$, *Range* = -19 – 22), where a negative score indicates dominance in English and a positive score indicates dominance in French. Sixteen parents were dominant in English, 25 were dominant in French, and 1 was equally dominant in both languages. In sum, while parents generally reported high levels of proficiency in both languages, most also reported having a dominant language. This reflects variation that is common between bilinguals. While this variation could explain parents' code-switching, such questions are beyond the scope of the current paper.

3.2.2 Procedure

Data for the corpus were collected as part of a larger research project on early bilingual development (Orena et al., 2019, 2020). The data for 10-month-olds were collected between November 8th, 2016 and September 18th, 2017, and the data for 18-month-olds were collected between July 25th, 2017 and March 28th, 2018. The audio recordings were collected using

Language ENvironment Analysis (LENA) devices, which are small, portable recorders that can record up to 16 hours. When infants were 10 months old, each family completed two appointments. At the first appointment, the procedure and purpose of the study were explained, and families were interviewed about their language use (LEQ via MAPLE; Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., 2020). Each family was given three LENA recording devices, and three infant vests to hold the devices. Families were asked to record three full days at home: two weekdays and one weekend day. Two families were unable to follow this schedule: one recorded 1 weekday and 2 weekend days, and the other recorded 3 weekdays. Three infants were enrolled in daycare at the time of their participation, but the recordings were made on days the infant was at home. Families were instructed to begin the recording when the infant woke up and have the LENA device record the entire 16 hours. When all three recordings were complete, there was another appointment where the LENA devices were collected and questionnaires about the parents' language experience and proficiency (LEAP-Q; Marian et al., 2007) and language mixing (Language Mixing Questionnaire; Byers-Heinlein, 2013) were administered.

When infants were 18 months old, families repeated the same procedure. To capture any changes in the infant's language environment, the Language Exposure Questionnaire and the Language Mixing Questionnaire were re-administered. At this age, for practical reasons families were only asked to record one weekend day. Most of the parents had finished their parental leaves, so most of the children were enrolled in daycare. Asking families to record their child's environment while they were in daycare during the week was not feasible, due to privacy concerns related to the presence of other children. For these reasons, families were asked to record one weekend day instead, in hopes that this would maximize participation.

3.2.3 Transcription

The LENA system does not differentiate between languages, and therefore cannot identify when code-switching occurs. Thus, language identification and transcriptions were conducted manually by trained, highly proficient, simultaneous French–English bilingual research assistants (for details, see Orena et al., 2020). To create the corpus, the recordings were first divided into 30-second segments, following a standard practice for coding daylong recordings (e.g., Ramírez-Esparza et al., 2014), and to allow research assistants to reliably pay attention to who was speaking and in what language (Orena et al., 2020). If an utterance broke off in mid-stream at the end of the segment, the research assistants listened to the following segment to transcribe the end of the utterance. Through pilot analyses of the corpus, it was determined that looking at every other segment was sufficient for evaluating infants' language environment (Orena et al., 2020). Therefore, the research assistants listened to every other segment and noted who was speaking, to whom, and in what language (see Figure 3.1). If any speaker used more than one language within a segment, the language of that segment was tagged as "mixed." Each of the mixed-language segments was transcribed by research assistants. The transcriptions were reviewed by a second group of research assistants to ensure accuracy.

3.2.4 Coding

Once all the segments that contained mixed language were identified and transcribed in the coded portion of the corpus, every instance where a parent was talking to their infant and changed the language they were speaking was tagged as a code-switch. This means that it was possible for a single segment to contain multiple code-switches if the speaker changed languages multiple times. If a single segment contained more than one code-switch, it was marked as such. The full coding protocol can be found at https://osf.io/yz6f7/ (see Appendix A).

Frequency

After identifying all instances of code-switching in the corpus, frequency was determined by normalizing this value by the amount of speech that children heard. This was important to ensure that observed differences in code-switching frequency would not simply reflect the overall level of interaction between infants and their caregivers. We used two related approaches to normalize our data. Our pre-registered approach was to calculate the number of code-switches per hour of speech directed at the infant, which was based on the number of 30-second segments that contained infant-directed speech. Our second approach, which was suggested by a reviewer and thus was not pre-registered, was to calculate the number of code-switches per 1,000 infant-directed words. This latter value was calculated based on LENA's automatic word count values, which have been shown to be reliable in this corpus (Orena et al., 2020). We were then able to compare these measures of frequency to parent-reported rates of code-switching from the Language Mixing Questionnaire. If the average speech rate that infants encounter is reasonably consistent, these two normalization approaches will yield similar results.

Direction

First, for each code-switch, we noted the direction of the switch. That is, we noted whether the speaker's language switched from French to English, or English to French. As there were no hypotheses related to this variable, it was not analyzed.

Syntactic location

Second, we determined the syntactic location of the switch. There were three possible levels for this coding. We determined whether the switch was between sentences (intersentential) or within a sentence (intrasentential). Any switch that happened between sentences and within the same 30-second segment was coded as intersentential, regardless of any intervening silence. For the intrasentential code-switches, we further determined whether the switch occurred within a syntactic phrase (e.g. the red *chien* [fr. dog]) or between syntactic phrases (e.g. *le chien* [fr. the dog] runs). This was determined by applying various tests of constituency (Radford, 2006).

Apparent reason for the code-switch

Lastly, each switch was coded for the apparent reason for the switch based on the context available in the audio segment. Apparent reasons were initially based on those previously reported in the literature on bilingual parents' code switching (Bail et al., 2015; Byers-Heinlein, 2013; Goodz, 1989): attracting the child's attention, adding emphasis, disciplining the child, bolstering the infant's understanding, attempting to teach new vocabulary, providing a translation, and conventionalized borrowings and phrases, including baby-specific words and phrases. For definitions for each of these reasons, see the coding manual at https://osf.io/yz6f7/ (see Appendix A). For examples, see Table 3.2. Additionally, after pilot coding 2 of the 10-month-old infants but before pre-registering the study, we decided to add baby-specific words and phrases as a subset of the borrowing category in order to better understand the nature of borrowing as a type of parental code-switching. Although we coded 8 different apparent reasons for code-switching drawn both from the literature and from our pilot coding, we acknowledge that our list is not exhaustive and that parents may code-switch for reasons not included here. Therefore, if a single switch did not appear to be motivated by any of our pre-determined reasons, the coding for the switch was left blank (i.e., categorized as "no reason"). Given the complex nature of code-switching and potential overlap between our categories, we allowed a single switch to be coded as having multiple apparent reasons.

3.2.5 Inter-Rater Reliability

To evaluate the accuracy of the data coding, inter-rater reliability was calculated for the following categories: direction, syntactic location, and each of the 8 apparent reasons for the code-switch. Data were initially coded by the second author, who has training in linguistics and psychology and is highly proficient in both French and English. Subsequently, the first author coded a randomly selected 20% of utterances to each infant. Inter-rater reliability for each category is reported as the percentage of code-switches for each category that were coded the same by both raters. Interrater reliability was generally high: 100% for the direction, 94% for the constituency of the code-switch, and ranging from 78% to 97% for each of the 8 apparent reasons (see Table 3.1). We pre-registered a minimum inter-rater reliability of 75% for each apparent reason category to be included in our analyses, thus all the categories were included in the subsequent analyses.

3.3 Results

All analyses were conducted as per our preregistration, except where deviations are noted. One important deviation is that we had originally planned to collapse the data across the two age groups (10 and 18 months old) for several of our analyses. However, after preliminary analyses revealed striking changes in parental code-switching across development, instead of reporting analyses that collapsed across age groups, we opted to report analyses for each age group separately followed by the planned statistical comparisons of the two ages. Coded data and analysis code are available at https://osf.io/bxkg7/.

3.3.1 Frequency

As a reminder, we operationalized frequency in two ways: number of code-switches per hour of infant-directed speech, and number of code-switches per 1,000 infant-directed words. To calculate the number of code-switches parents produced per hour of infant-directed speech, we divided the number of parental code-switches by the number of 30-second segments where a parent was speaking to their infant and then multiplied this number by 120, the number of segments per hour. To calculate the number of code-switches per 1,000 infant-directed words, we divided the number of parental code-switches by the number of infant-directed words (as calculated by LENA's automated word counter). At 10 months, because there were three days of recording for each child, we averaged the frequency of code-switching across the three days. The two measures of frequency were highly correlated, $r = 0.91$, 95% CI [0.82, 0.95], $t(35) = 12.605$,

$p < 0.001$, and thus results were highly similar whether calculated by hour of infant-directed speech or per 1,000 infant-directed words.

On average, 10-month-olds heard 7 ($SD = 4.85$, *Range* = 1 – 16) code-switches per hour of infant-directed speech, which corresponded to 6 code-switches per 1,000 infant-directed words ($SD = 3.91$, *Range* = 0 – 13). On average, 18-month-olds heard 28 ($SD = 22.03$, *Range* = 1 – 84) code-switches per hour of infant-directed speech, which corresponded to 18 code-switches per 1,000 infant-directed words ($SD = 16.21$, *Range* = 0 – 59). Paired *t*-tests of the families that contributed recordings at both ages confirmed that parents code-switched more frequently in interactions with 18-month-olds than in interactions with 10-month-olds, whether measured by code-switches per hour of infant-directed speech $t(15)=-3.89$, $p=.001$, $M_d=-19.99$, 95% CI [-30.94, -9.03] (see Figure 3.2a) or code-switches per 1,000 infant-directed words $t(15) = -3.26$, $p = .005$, $M_d = -11.71$, 95% CI [-19.36, -4.06] (see Figure 3.2b).

In an exploratory analysis, we examined how the frequency of code-switching may have changed in each individual family. We compared each family's code-switching at 18 months to their code-switching at 10 months. If a family's code-switching at 18 months increased or decreased from their code-switching at 10 months by more than 2 times the full sample's standard deviation at 10 months (per hour of infant-directed speech: $SD = 4.85$; per 1,000 words: $SD = 3.91$), we considered this to indicate a change in the frequency of code-switching within a family. Our rationale was that a change in frequency of less than 2 standard deviations could be attributed to normal variation within the range of what had been observed across families at 10 months, but a change greater than 2 standard deviations would indicate a meaningful difference. As measured by the number of code-switches per hour of infant-directed speech, 9 families increased the frequency of their code-switching, 7 families did not change the frequency of their code-switching, and no families decreased the frequency of their code-switching. As measured by the number of code-switches per 1,000 infant-directed words, 8 families increased the frequency of their code-switching, 8 families did not change the frequency of their code-switching, and no families decreased the frequency of their code-switching. Thus, code-switching appeared to generally remain stable or increase across these two time points.

3.3.2 Validity of the Language Mixing Questionnaire

To evaluate the validity of the Language Mixing Questionnaire, parents' responses to the questionnaire were compared to their code-switching behaviors observed in the data. To do this, each parent who completed the questionnaire was assigned a Language Mixing Scale Score (following Byers-Heinlein, 2013), calculated by summing the responses to 5 questions on the questionnaire with Likert scales (1 = very true, frequent language mixing; 7 = not at all true, infrequent language mixing). This sum was then subtracted from 35, the highest possible sum. This resulted in a maximum score of 30 for those who report frequently code-switching, and a minimum score of 0 for those who report never code-switching. One parent did not have a Language Mixing Scale Score and was excluded from the following analyses. Parents had a mean Language Mixing Scale Score of 11.07 (SD = 8.73; *Range* = 0 - 30) at 10 months and 12.48 (SD = 7.59; *Range* = 1 - 28) at 18 months.

In our pre-registration, we had planned to compute a correlation between parents' Language Mixing Scale Score and a quantification of code-switching frequency where we would divide the number of code-switches each parent produced by the number of 30-second segments where they spoke to their infant and multiplying this by 100. However, we deviated slightly from this plan, to instead be consistent with the quantifications of code switching used in the previous analyses: the number of code-switches each parent produced per hour of infant-directed speech, and the number of code-switches each parent produced per 1,000 infant-directed words. We note that the metric of code-switches per hour is a linear transformation of our pre-registered metric of code switches per 30 seconds * 100, and thus this change does not impact inferential statistics.

Because the Language Mixing Questionnaire was administered at both 10 and 18 months, it was possible to compute correlations between self-reported and observed code-switches at two ages, and thus scores from each age were included as separate data points in the following analyses. The correlation between the Language Mixing Scale Scores and parents' observed code-switching was statistically significant, with a moderate effect size, for both the number of code-switches per hour of speech, r = .37, 95% CI [.15, .56], $t(69)$ = 3.30, p =.002 (see Figure 3.3a), and per 1,000 infant-directed words, r = .35, 95% CI [.13, .54], $t(69)$ = 3.13, p = .003 (see Figure 3.3b).

One previous study found that the Language Mixing Scale Score has a higher correlation with parents' actual intersentential code-switching than intrasentential code-switching, despite the questionnaire asking mainly about intrasentential code-switching (Bail et al., 2015). To examine the replicability of this finding, we conducted additional analyses that considered intersentential and intrasentential code-switches separately. Parents who participated at both ages have different intersentential and intrasentential frequencies for each age. The correlation between the frequency of intersentential code-switching and the Language Mixing Scale Scores was statistically significant for both the number of code-switches per hour of speech, $r = .34$, 95% CI [.11, .54], $t(67) = 2.98$, $p = .004$, and per 1,000 infant-directed words, $r = .33$, 95% CI [.10, .52], $t(67) = 2.84$, $p = .006$. The correlation between the frequency of intrasentential code-switching and the Language Mixing Scale Scores was of a similar magnitude and direction, and was statistically significant for both the number of code-switches per hour of speech, $r = .30$, 95% CI [.07, .50], t(67) = 2.61, $p = .011$, and per 1,000 infant-directed words, $r = .29$, 95% CI [.05, .49], $t(67) = 2.45$, $p = .017$.

To compare the correlations between the Language Mixing Scale Scores and the intersentential and intrasentential frequencies directly for each frequency measure, we transformed them using Fisher's r to z transformation. Comparing these dependent, overlapping correlations revealed that the correlations between the intersentential and intrasentential frequencies and the Language Mixing Scale Score were not statistically significantly different for either the number of code-switches per hour of speech, $z = 0.38$, $p = 0.70$, or per 1,000 infant-directed words, $z = 0.42$, $p = 0.68$. Additionally, parents' intersentential and intrasentential frequencies were correlated for both the number of code-switches per hour of speech, $r = .60$, 95% CI [.43, .73], $t(68) = 6.20$, $p < .001$, and per 1,000 infant-directed words, $r = .62$, 95% CI [.45, .75], $t(68) = 6.53$, $p < .001$, suggesting that parents who code-switch intersententially also code-switch intrasententially, which could explain why parents' Language Mixing Scale Score was similarly correlated with both types of directly-observed code-switching.

3.3.3 Syntactic Location

3.3.3.1 Frequency comparison of intersentential and intrasentential code-switching

To evaluate our prediction that parents would produce more code-switches between sentences than within a sentence, we divided the number of intersentential code-switches by the total number of code-switches at each age of recording. An intersentential percentage score of 50% would therefore indicate that intersentential and intrasentential code-switches happened at the same rate. At 10 months, on average, 77% (*Range* = 50% – 100%) of code-switches were intersentential. At 18 months, on average, 83% (*Range* = 61% – 100%) of code-switches were intersentential. We conducted a one-sample *t*-test with $\mu_0 = 50$ at each age. Consistent with our predictions, parents produced more intersentential (e.g., Come on. *C'est fini.* [It's done.]) than intrasentential (e.g., *Est-ce qu'on va aller manger* [Are we going to eat] banana pancake?) code-switches at both 10 months, $t(20) = 7.85$, $p < .001$, $M = 77.11$, 95% CI [69.91, 84.32], and 18 months, $t(15) = 11.73$, $p < .001$, $M = 82.90$, 95% CI [76.92, 88.88].

Next, we examined whether the percentage of intersentential code-switches changed across development. A paired *t*-test for the 16 families that provided recordings at both ages revealed that parents code-switched intersententially more when their child was 18 months old (83%) than 10 months old (74%), $t(15) = -2.21$, $p = .043$, $M_d = -8.47$, 95% CI [-16.64, -0.29]. The change in the percentage of code-switches at each syntactic location across ages can be seen in Figure 3.4.

3.3.3.2 Frequency comparison of intrasentential code-switching at and within syntactic boundaries

To evaluate our prediction that within-sentence code-switches are more likely to occur between syntactic phrases than within syntactic phrases, we divided the number of intrasentential code-switches that occurred between syntactic phrases by all code-switches that occurred within a sentence. A between-phrase percentage score of 50% would therefore indicate that intrasentential code-switches between and within syntactic phrases happen at the same rate. At 10 months, on average, 62% (*Range* = 0% – 100%) of intrasentential code-switches occurred at a syntactic boundary. At 18 months, on average, 54% (*Range* = 14% – 100%) of intrasentential code-switches occurred at a syntactic boundary. We conducted a one-sample *t*-test with $\mu_0 = 50$ at each age to examine if the percentages of intrasentential code-switches produced between and within syntactic phrases were equivalent. These tests revealed that the intrasentential percentage score was not statistically significantly different from 50% at either 10 months, $t(17) = 1.54$, $p = .143$, $M = 61.83$, 95% CI [45.58, 78.08], or 18 months, $t(14) = 0.73$, $p = .480$, $M = 54.46$, 95% CI [41.28, 67.64]. Our results did not support the prediction that parents produce more intrasentential code-switches at a syntactic boundary (e.g., Now you want *lait* [milk].) than within a syntactic phrase (e.g., *C'est un* [It's a] monkey.).

Next, we examined whether the percentage of intrasentential code-switches changed across development. A paired sample *t*-test of the families that provided recordings and produced intrasentential code-switches at both ages revealed that there was no statistical difference in the rate of between-phrase percentage scores across time points $t(13) = 0.65$, $p = .529$, $M_d = 6.86$, 95% CI [-16.06, 29.78]. This indicates that while the frequency of code-switching increases between 10 and 18 months, the percentage of intrasentential code-switches occurring at and within syntactic boundaries remains stable.

3.3.4 Apparent Reason

3.3.4.1 Co-occurrence of apparent reasons

Because our coding system allowed for a single code-switch to be coded as having multiple apparent reasons, we wanted to evaluate if two reasons co-occurred frequently enough to be combined into a single reason. Thus, for each of the 8 reasons, we calculated the proportion of switches coded for that particular reason that were also coded for each of the other 7 reasons. We identified two pairs of reasons with a co-occurrence rate above 75%, a value set in our pre-registration. First, 100% of the code-switches that were attributed to the use of baby-specific words were also coded as language borrowing. This was unsurprising given that the baby word category was added as a subset of borrowing. Second, 80% of the code-switches that were attributed to the use of translation equivalents were also coded as increasing understanding. Following our pre-registration, we combined each pair of reasons that frequently co-occurred into a single category. Additionally, we kept each of the original reasons as subsets of the combined category for subsequent analyses.

3.3.4.2 Frequency of apparent reasons

To explore the frequency of each apparent reason, we calculated the proportion of code-switches motivated by that reason for each parent. The proportions for each reason were then averaged across all parents. We created a contingency table with the time points and apparent reasons as factors (see Table 3.2). No statistical tests were planned or conducted, as we had no specific prediction regarding the frequency of the different apparent reasons.

Parents appear to code-switch most frequently in an effort to bolster their child's understanding. Moreover, while common borrowings of words and phrases were relatively frequent in our data, these borrowings did not appear to be attributable to the use of baby-specific words or phrases. The most notable change across time points was the increase in teaching vocabulary. Other apparent reasons were not frequent in our data but do seem to motivate some of the parents' code-switching. Finally, we observed very few code switches that did not seem to fit any of the apparent reasons we coded, indicating that most parental code-switches fit into one or more of these categories.

The frequency of each of these reasons motivating a code-switch also varied across parents. Figure 3.5 shows the percentage of code-switches that were attributed to parents 1) bolstering their child's understanding and/or producing a translation equivalent, 2) bolstering understanding and one of the other 6 apparent reasons, and 3) only another apparent reason. These mutually exclusive categories were created to illustrate the prevalence of understanding as an apparent reason for code-switching relative to the other reasons. The numbers in each bar represent the count for each of the three categories. This figure shows not only the variability in apparent reasons behind parents' code-switching, but also the variation in frequency of code-switching by individual parents.

3.4 Discussion

In this study, we investigated the properties of parents' code-switching behaviors in everyday interactions with their infant. Specifically, we used a corpus of at-home recordings to analyze how frequently French–English bilingual parents in Montreal code-switched, as well as the syntactic location and apparent reason for each of their code-switches. First, we found that the frequency of code-switching, whether controlling for hours of infant-directed speech or number of infant-directed words, generally increased between 10 and 18 months of age. Second, we found that the majority of parents' code-switches occurred intersententially at both ages, and that this proportion increased across their infant's development. For the code-switches that occurred intrasententially, the proportions of code-switches that happened between syntactic phrases and within a syntactic phrase were comparable at both ages. Last, while parents code-switched for a variety of apparent reasons, most parental code-switches at both time points appeared to be motivated by the desire to bolster their infant's understanding. Parents also appeared to code-switch more to teach vocabulary when their infant was 18-months old than when they were 10-months old. Combined, our results suggest that parents may be adapting their code-switching behavior to their infant's developing linguistic abilities, producing code-switching that could support successful acquisition of both languages.

The first indication that parents may be adapting their code-switching to their infant's language abilities is the increased frequency of code-switching between 10 and 18 months. Between these two ages, an infant's language abilities undergo a large transformation: at 10 months, most infants do not produce a single word, whereas at 18 months, infants may be producing as many as several dozen (Fenson et al., 2014). It is possible that at 18 months, parents are aware of which words an infant knows and which language those words are in. Parents may then code-switch more to strategically support their infant's language development in two different ways. One way that parents may code-switch strategically is by switching languages to use a word they believe their infant understands. This pattern is consistent with the current data showing that parents produced a higher total number of code-switches to bolster their infant's understanding and/or to provide a translation equivalent when their infant was 18 months old compared to when they were 10 months old.

Another way parents may code-switch strategically is by switching languages to use a word they believe their infant does not understand in order to teach them a new word. This pattern is also consistent with our data, as parents were found to code-switch to teach vocabulary more when their infant was 18 months old compared to when they were 10 months old. For example, this could explain the positive relationship between parents' intrasentential code-switching and their child's vocabulary size found in previous research (Bail et al., 2015). While these two reasons for code-switching are seemingly paradoxical, in conjunction, they could ultimately support the acquisition of two languages.

Parents may also adapt their code-switching to their infant's language abilities through altering the syntactic location of their code-switches. Consistent with previous research, at both time points, the majority of code-switches that the parents produced occurred intersententially (Bail et al., 2015). Parents may use more intersentential code-switches when speaking to their infant because intersentential code-switches are easier to produce than intrasentential code-switches (Poplack, 1980). The relative difficulty speakers have in producing intrasentential code-switches is mirrored by processing difficulties for listeners in comprehending them. Experimental work has suggested that intrasentential, but not intersentential, code-switches elicit processing costs in bilingual infants (Byers-Heinlein et al., 2017; Potter et al., 2018), and thus the majority of code-switches that bilingual parents produce are those that are the least difficult for their infants to understand. The processing costs associated with intrasentential code-switches may underlie parents' shift toward producing a higher percentage of intersentential code-switches at 18 months compared to 10 months. Parents may (likely implicitly) realize that intrasentential code-switches are difficult for their infant to understand, so they decrease the number of intrasentential code-switches they produce to reduce processing costs, thus supporting their infant's comprehension and resulting in a higher percentage of intersentential code-switches. It is unlikely that parents produced more intersentential code-switches when their child was 18 months old simply because they are easier to produce. If a parent is able to produce intrasentential code-switches when their infant is 10 months old, they likely retain that ability eight months later when their infant is 18 months old. Therefore, any changes in the production of code-switching are probably due to external influences, in this case, the development of their infant. If parents are indeed altering their code-switching behavior in an effort to reduce processing costs for their infant, this suggests that aspects of parental speech unique to bilingual contexts are sensitive to an infant's linguistic development.

One prediction that was not supported by our analyses was that parents' intrasentential code-switches would occur more often at a syntactic boundary than within a syntactic phrase. Instead, we found that these occurred at a similar frequency. This result may be driven by single-word code-switches that occur between a determiner and a noun (e.g., the *chien* [dog]), which has been found to be a frequent location for parental code-switching (Bail et al., 2015). However, our coding scheme did not record the exact syntactic location or the number of words that followed a given code-switch, which could be addressed in a subsequent study. Therefore, it is possible the frequency of single-word code-switches could explain the equivalent proportion of intrasentential code-switches at and within syntactic boundaries. Future work is needed to confirm this prediction.

In sum, our results suggest that, similar to other aspects of infant-directed speech, infant-directed code-switching can have qualities that might support infant language development. Future naturalistic studies could examine links between parents' use of supportive code-switching strategies, and infants' language outcomes. In addition, laboratory studies could directly investigate whether code-switching supports bilingual infants in learning words in each of their languages. By describing the quality and quantity of code-switching that children hear, we can ask more nuanced questions about how code-switching affects bilingual language development.

3.4.1 Differences in Parental Code-Switching Patterns Between Bilingual Communities

The current study focused on parental code-switching patterns in one bilingual community: French–English bilingual families in Montreal. Given the limited research on parents' naturally produced code-switching, it is unknown how much these patterns generalize to other bilingual communities. Understanding the differences between bilingual communities could be important when synthesizing findings on bilingual language development. Parental code-switching in different bilingual communities may have different properties, which may impact language development in different ways, such as the potential link between frequency and a child's vocabulary size (Bail et al., 2015; Byers-Heinlein, 2013; Place & Hoff, 2016).

To illustrate, here we compare our findings to Spanish–English bilingual parents' code-switching during a laboratory play session with their 17- to 24-month-olds, the only other study to our knowledge to directly investigate and describe parents' code-switching (Bail et al., 2015). It is important to note that the majority of parents in both our study and this study reported being highly proficient in both of their languages. By comparing the results of these two studies, one major difference between French–English parents in Montreal and Spanish–English parents in the U.S. stands out: the frequency of parental code-switching. Spanish–English parents code-switched, on average, more than 30 times in a 13-minute play session – over four times more than French–English parents, who code-switched, on average, 28 times in an hour of speech when their child was 18 months old.

One highly plausible explanation for this difference is that different bilingual communities may have different baseline rates of code-switching that permeate into parents' code-switching with their children. Code-switching may simply be more frequent in Spanish–English communities in the U.S. compared to French–English communities in Canada. While it is hard to determine the exact underlying cause(s) of the difference in code-switching frequency across communities, it is possible that communities use code-switching in different ways to create and maintain a group identity (Nilep, 2006). In Canada, French and English are both official languages, and in Montreal, both languages are widely used throughout the community and have high sociolinguistic status. Given the prevalence of both languages in the larger community, code-switching may not be used by French–English bilinguals to maintain a group identity (Kircher, 2009). However, bilingual Spanish–English communities in the U.S. may feel more of a need to cultivate a group identity through the use of frequent code-switching, due to the minority status of Spanish in the larger community (Zentella, 1981).

A second, complementary possibility is that observed differences between these studies are attributable to divergent methodologies, with different ages of participants, procedures, and coding approaches. First, the age of a child may be an important factor in influencing how parents code-switch. For example, the Spanish–English parents might have code-switched more than the French–English parents, because the Spanish–English sample was older (ranging from 17 to 24 months old). Given our results suggesting that the frequency of code-switching increases across a child's development, it would not be surprising that the older Spanish–English children heard more code-switching than the younger French–English children.

Second, the differences between communities may be explained by the methods used to collect the speech samples. Short, structured play sessions result in denser speech samples and different features of speech (e.g., density of noun input) compared to naturalistic at-home recordings (Belsky, 1980; Bergelson et al., 2019; Tamis-LeMonda et al., 2017). If parents' speech is denser in play sessions, this provides more opportunities for them to code-switch. However, we were able to control for the density of speech in our analysis in the number of code-switches parents produced per 1,000 words. Therefore, it is unlikely that the speech density between play sessions and at-home recordings underlie the differences in the rate of code-switching between the two communities. Additionally, parents might code-switch at a different frequency during daily life as compared to play sessions, particularly in the lab. Therefore, the frequency of code-switching in play sessions may be inflated compared to the frequency of code-switching in daily life. Other methodological differences, such as the number and gender of parents included in the sample could also contribute to the different findings. Therefore, differences between Spanish–English and French–English communities might be attenuated if parents' code-switching was assessed using the same method.

Lastly, the way in which the recordings were made and transcribed in the current study could be underestimating how frequently French–English parents code-switch. First, our transcription was only able to capture code-switches that happened within the same 30-second segment. This method could have missed code-switches that happened between segments. However, it is unlikely that enough code-switches occurred at these precise boundaries to dramatically alter our results. Second, we only coded every other segment. While pilot analyses determined that this resulted in a sample that sufficiently represented an infant's language environment (Orena et al., 2019), some of these segments may have had higher levels of code-switching. Lastly, only one weekend day was recorded at 18 months, compared to the 2 weekdays and 1 weekend day at 10 months. This may not have captured the child's entire linguistic environment, therefore, our estimate of code-switching frequency in Montreal may not be fully representative for older infants. Future studies applying the same methods will be needed to directly address the question of how code-switching varies across bilingual communities.

3.4.2 Methodological Contributions

Beyond the substantive contributions toward understanding the nature of parents' code-switching, this study provided several methodological contributions. To our knowledge, this study was the first to investigate parental code-switching at home through daylong recordings, and we were able to develop several novel approaches to do so. An important feature of our coding scheme is the ability to measure the frequency of different apparent reasons that parents code-switch throughout their daily life. Determining why code-switches occur is difficult even for the speaker producing the code-switch (Gumperz, 1982), so at the outset it was unclear whether this could be reliably coded. However, the main coder carefully considered the context of each switch when assigning the apparent reason(s) for the switch and the interrater reliability for each of the individual reasons was high. Additionally, fewer than 1% of the code-switches in the dataset were not coded as being motivated by any of our predetermined reasons, suggesting that the reasons we examined are representative of why parents code-switch when speaking to their child. This suggests that our approach can reasonably determine the apparent reason behind a parent's code-switch.

Second, we were able to assess the relationship between parents' actual code-switching frequency and their responses to the Language Mixing Questionnaire (Byers-Heinlein, 2013). These two measures were found to have a statistically reliable correlation, (r = .30 – .34), suggesting that the Language Mixing Questionnaire can detect some of the variation in the frequency of parents' code-switching. Nonetheless, the Language Mixing Scale Scores only explained 14% of the variance in parents' code-switching frequency. There are several possible explanations for this result. First, parents may be unable to answer the questions on the questionnaire accurately, due to a lack of awareness of their use of code-switching or not understanding what the questionnaire is asking (Myers-Scotton, 2017). Second, the range of code-switching observed in the data was restricted, particularly at 10 months. It is a well-known statistical phenomenon that the magnitude of a correlation is reduced when a sample has a restricted range of scores. The Language Mixing Questionnaire may not be fine-grained enough to pick up on variation in the frequency of code-switching when it is relatively infrequent, as it was in our data.

3.4.3 Scope and Future Directions

While this study provides the first account of parents' naturally produced code-switching, we nonetheless had to limit our scope to what could be reasonably explored in one study. There are still many other questions that this and similar datasets could address in future research. There are two major directions we propose for this research: investigating predictors of parental code-switching and investigating how parental code-switching may be linked to child language outcomes.

First, we did not explore whether demographic variables (e.g., parental language proficiency or dominance, familial language strategy) impacted parents' code-switching behaviors. Research has been able to identify some predictors of an adult's rate of code-switching when speaking to another adult, such as personality and language history (Dewaele & Li, 2014). This research has not yet been extended to when adults are speaking to children.

Second, we also did not investigate impacts of parents' code-switching on infants' linguistic development, such as vocabulary scores. Our focus was on investigating the variation in a bilingual infant's environment, which we believe lays crucial groundwork for understanding how this variation affects infants' language development. This is an important direction for future research, because there is little consensus in the literature on whether parents' code-switching affects their infant's language development (Bail et al., 2015; Byers-Heinlein, 2013; Orena et al., 2019; Place & Hoff, 2016). It is possible that the inconsistent findings are due to qualitative and quantitative differences in the code-switching parents produce across different bilingual communities. Thus, more research, applying the same or similar methods as used in this study to different bilingual populations, is required to strengthen this foundational understanding of how infant-directed code-switching varies across communities. Once community differences are better understood, future research could then build upon this knowledge and examine the direct impact of parents' code-switching on children's language development or bilingual language development in general.

3.5 Conclusion

Code-switching is a linguistic phenomenon that is pervasive in bilingual and multilingual communities; thus, it is unsurprising that bilingual parents code-switch when speaking to their infants. Our results from a sample of French–English bilingual families in Montreal show that the frequency of parents' code-switching and the percentage of intersentential code-switches increased between 10 and 18 months of age. At both ages, parents appeared to code-switch most frequently in order to boost their child's understanding. At 18 months, parents code-switched to teach vocabulary more than they had when their infant was 10 months old. Combined, these results suggest that parents may code-switch in ways that support successful bilingual language development.

Table 3.1: Inter-rater reliability for each apparent reason.

Category	Percent Agreement
Attention	91
Baby words	98
Borrowing	92
Discipline	97
Emphasis	85
Translation equivalent	85
Understanding	78
Vocabulary	91

Table 3.2: Percentage (raw count in parentheses) of code-switches observed for each apparent reason at 10 and 18 months, the difference in percentage (difference in raw count in parentheses) across ages, and examples of each reason.

Reason	10 Months	18 Months	Difference	Examples
Understanding & Translation equivalent	74.2 (548)	74.7 (926)	0.5 (378)	1. *Papa travaille.* [Daddy's working.] Daddy's working, okay? 2. La lumière. [The light]. It's the light.
Understanding	73.0 (538)	73.6 (899)	0.6 (361)	1. I wouldn't eat that. *Pas pour manger.* [Not to eat] 2. One more? *C'est le dernier*. [It's the last one.]
Translation equivalent	7.7 (59)	6.1 (128)	-1.6 (69)	1. Hi. *Bonjour*. [Hello.] 2. Shark. *Requin*. [Shark.]
Borrowing	12.7 (90)	11.4 (100)	-1.3 (10)	1. It's *dodo* [nap] time. 2. *C'est* [That's] cool.
Non-baby words	11.6 (79)	10.4 (77)	-1.2 (-2)	1. Is it good? *Bon appétit*. [Enjoy your meal.] 2. Hey, if that's all it takes honey, *la vie est belle* [life is beautiful.]
Baby words	1.0 (11)	1.0 (23)	0.0 (12)	1. You want the *suce* [pacifier]? 2. Ya you have four *doudous* [blankies].
Emphasis	9.4 (71)	6.5 (135)	-2.9 (64)	1. A bear! *Oui*! [Yes!] 2. Gentle gentle. *Comme ça*. [Like that.]
Discipline	6.1 (34)	5.2 (30)	-0.9 (-4)	1. Come here. *Touche pas.* [Don't touch.] 2. Hey. *Fais pas ça*. [Don't do that.]
Vocabulary	3.5 (35)	8.3 (122)	4.8 (87)	1. *C'est noir.* [It's black.] And that's gold! 2. Can you say *gazon* [grass]?
Attention	3.5 (27)	1.2 (22)	-2.3 (-5)	1. Hi. *Regarde-moi*. [Look at me.] 2. Here sweetie. *Allô*. [Hello.]
No reason	0.4 (4)	0.3 (3)	-0.1 (-1)	1. *C'est* [That is], yeah? 2. Flyer. *P'tites choses.* [Little things.]

Note. A code-switch could be coded as having multiple apparent reasons.

Figure 3.1: Transcription and coding pipeline describing the number of coders and segments at each stage.

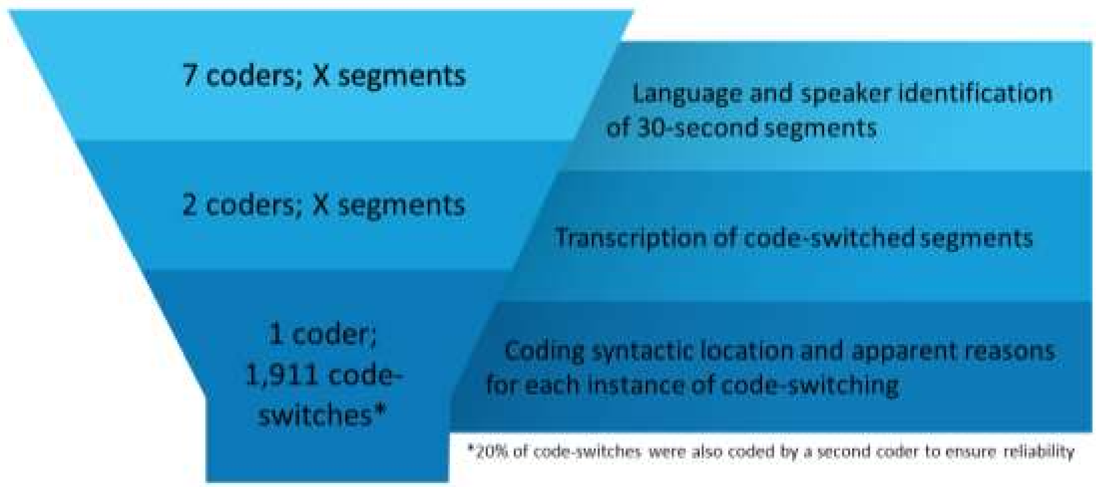
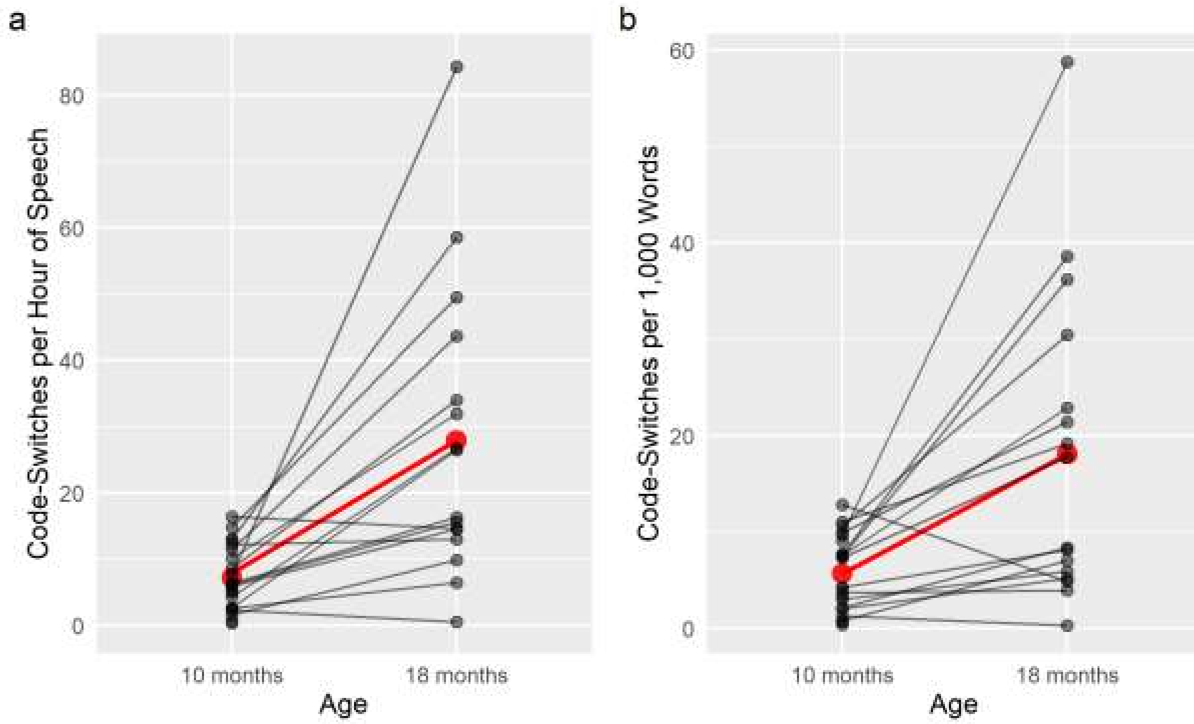


Figure 3.2: Change in the number of code-switches per hour of infant-directed speech for individual families between 10 and 18 months of age (a) per hour of infant-directed speech and (b) per 1,000 infant-directed words.



Note. The grey points and lines represent individual families, and the red points and line show the average change.

Figure 3.3: The relationship between parents' Language Mixing Scale Score and the number of code-switches per (a) hour of speech and (b) 1,000 infant-directed words based on data collected at 10 and 18 months combined.

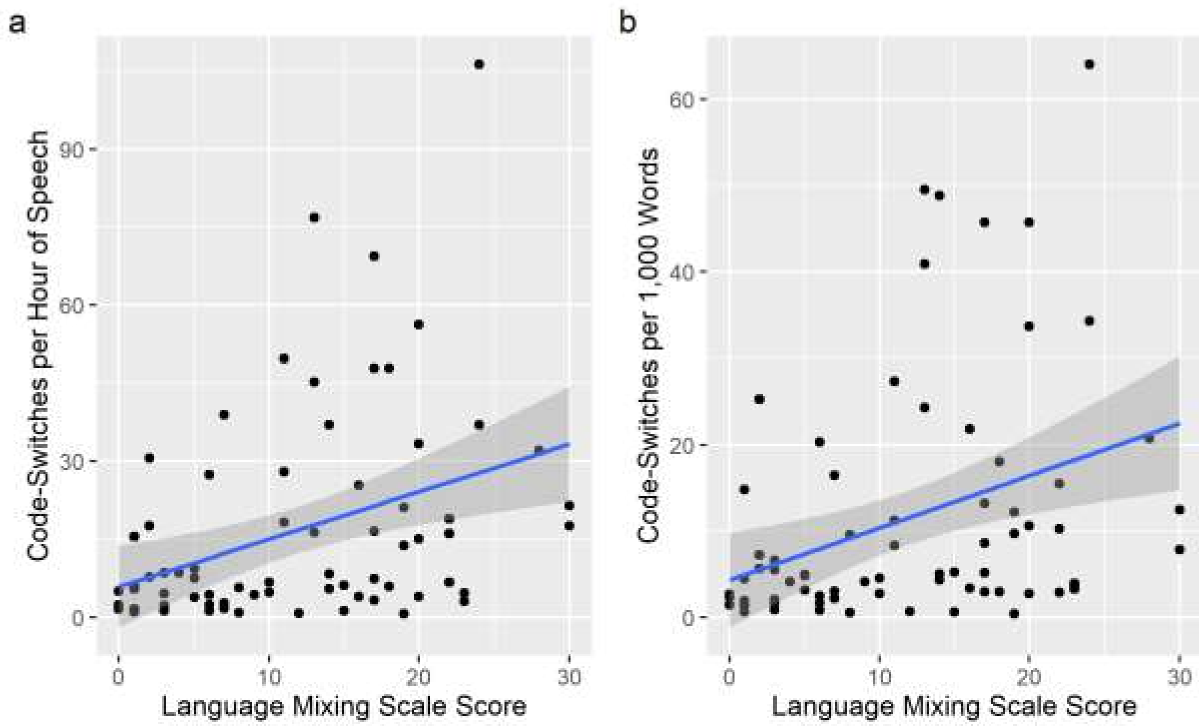


Figure 3.4: The percentage of code-switches produced at different syntactic locations across ages with all families included at each age.

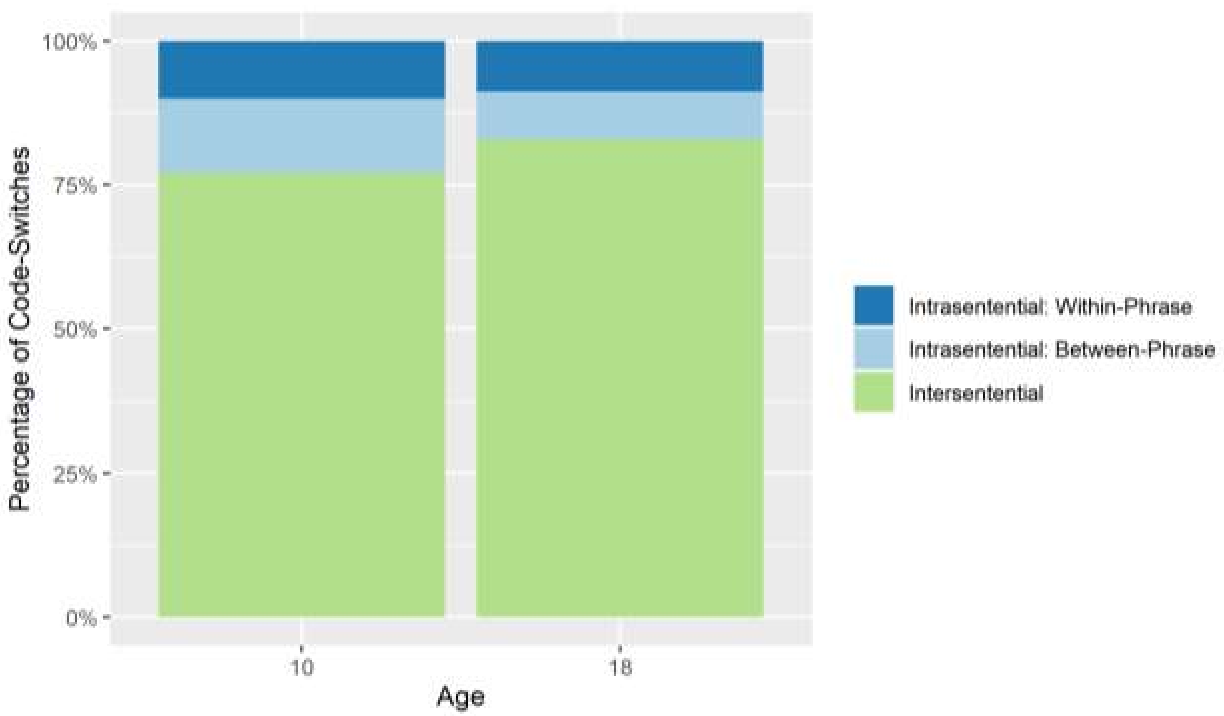
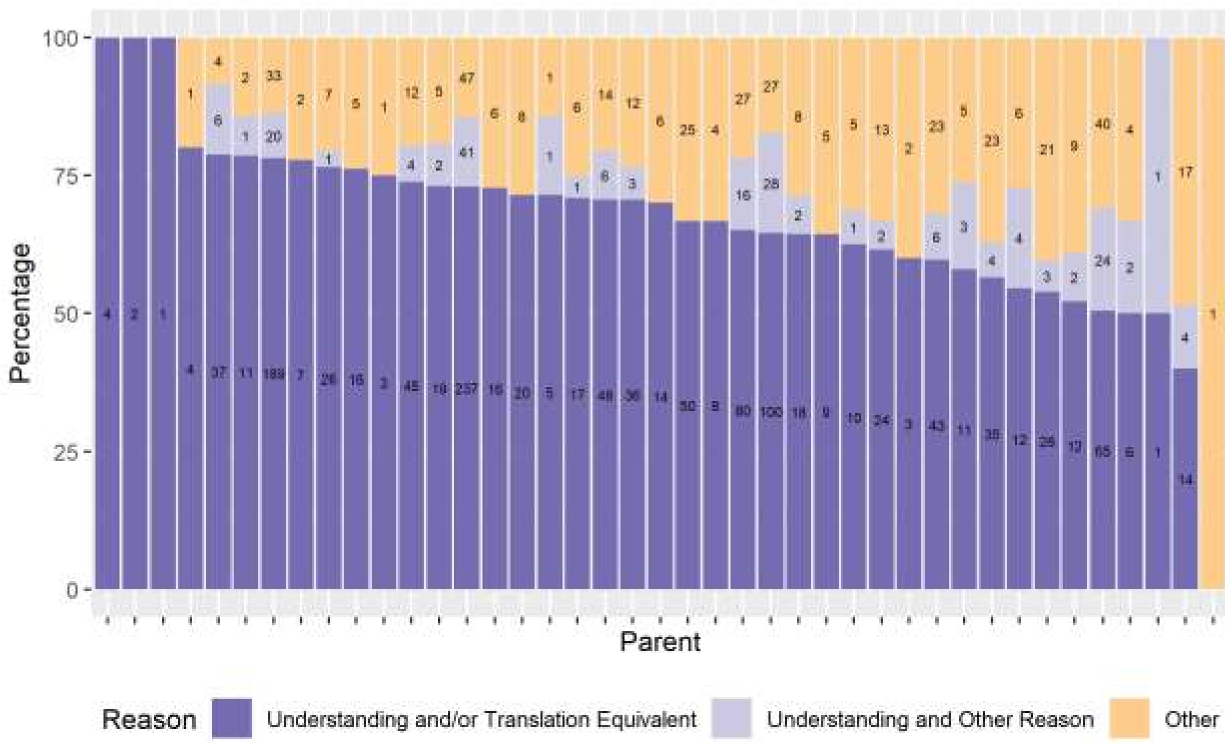


Figure 3.5: Percentage and count of parents' code-switches motivated by apparent reason.



4 Bilingual Children's Comprehension of Code-Switching at an Uninformative Adjective

4.1 Introduction

Bilingual children regularly hear both of their languages within a single conversation and even within a single sentence (e.g., *C'est un* [fr. It's a] monkey.). This phenomenon is known as code-switching. Most bilingual children hear code-switching in their daily lives (Chapter 3), and there is some evidence that over time code-switching may impact a child's vocabulary size (Bail et al., 2015; Byers-Heinlein, 2013) and overall language development (Kaushanskaya & Crespo, 2019). Code-switching can also reduce a child's comprehension in the moment as they process speech (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019). To date, research on children's comprehension of code-switching has focused on code-switches at a noun (e.g., "*Dónde está la* [sp. where's the] ball?"), even though everyday code-switching happens at many different parts of speech, such as verbs, prepositions, and adjectives (e.g., "*C'est* [fr. It is] yucky."; MacSwan, 2012). Here, we extend previous findings with nouns and investigate how code-switching at a mid-sentence determiner-adjective pair affects bilingual children's language comprehension.

A large body of literature has reported that bilingual adults process code-switches more slowly than single-language stimuli (for recent reviews see Beatty-Martínez et al., 2018; Valdés Kroff et al., 2018; van Hell et al., 2018), but researchers have only recently begun to study how young children process code-switches using looking-while-listening procedures. One eye-tracking study indicated that children process code-switches differently depending on whether the switch happens between sentences or within a single sentence. When hearing between-sentence code-switching (e.g., "That one looks fun! *Le chien* [fr. the dog]!"), 1.5- to 2-year-old children were as accurate at identifying the target object as they were when hearing a single language (e.g., "That one looks fun! The dog!"; Byers-Heinlein et al., 2017). However, when hearing within-sentence code-switching (e.g., "Look! Find the *chien* [fr. dog]!"), children were less accurate at identifying the target object compared to hearing a single language (e.g., "Look! Find the dog!"; Byers-Heinlein et al., 2017; Morini & Newman, 2019). Such studies with young children have focused solely on code-switches at the noun, so they do not address the potential impact of code-switching at other parts of speech. This limitation makes it impossible to draw generalized conclusions about how code-switching may or may not affect comprehension. Children may process code-switching at different parts of speech more readily depending on several factors, such as how often children hear code-switching in that location or what functional information is contained in the code-switched word(s). Evaluating children's comprehension of code-switching at different parts of speech will allow us to adjudicate between two general accounts of what makes code-switching difficult to process, which we describe as the frequency account and the functional account.

4.1.1 Frequency Account

The frequency account posits that how easily bilinguals process a code-switch depends on how frequently that type of code-switched construction occurs in their everyday life (e.g., Abutalebi et al., 2007; Guzzardo Tamargo et al., 2016). This account predicts that frequent code-switched constructions will be more easily processed than infrequent code-switched constructions. For example, in one study, Spanish–English bilingual adults more readily processed a common code-switch that included an entire compound verb (e.g., "*los senadores* [sp. the senators] <u>have requested</u> the funds") than an uncommon code-switch that occurred in the middle of the compound verb (e.g., "*los senadores <u>han</u>* [sp. the senators <u>have</u>] <u>requested</u> the funds"; Valdés Kroff et al., 2018). Similarly, Welsh–English bilingual adults judged code-switching at common parts of speech, such as nouns, to be more acceptable than code-switching at uncommon parts of speech, such as adjectives (Vaughan-Evans et al., 2020). The frequency account could also predict differences in comprehension between bilingual populations if they hear different rates of code-switching in their daily lives (Gosselin & Sabourin, 2021; Valdés Kroff et al., 2018).

If frequency is indeed an important factor in how bilingual adults process code-switching, its importance could also extend to children's processing. Under the frequency account, children would be expected to understand code-switching at frequently code-switched parts of speech, such as nouns, more easily than at infrequently code-switched parts of speech, such as adjectives. This account could explain existing findings about children's processing of code-switching. Children hear more between-sentence code-switches than within-sentence code-switches from their parents (Bail et al., 2015; Chapter 3), so the frequency account is consistent with the experimental finding that children more easily process between-sentence code-switches compared to within-sentence code-switches (Byers-Heinlein et al., 2017; Morini & Newman, 2019). When children do hear within-sentence code-switching, it often occurs at a noun (Bail et al., 2015). Thus, if within-sentence code-switches at a relatively common location for code-switching (i.e., the noun) disrupt children's processing, then within-sentence code-switches at an uncommon location should be even more disruptive.

4.1.2 Functional Account

The functional account proposes that bilinguals process code-switches differently based on the functional properties of the code-switched word(s), including grammatical properties. While prior research has investigated a variety of functions of code-switching in production – such as adding emphasis, signaling community identity, and facilitating understanding (Goodz, 1989; Heredia & Altarriba, 2001; Nilep, 2006) – comprehension studies have mainly focused on the functional dimension of grammatical class. One study of German–Russian bilingual adults used event-related potentials (ERPs) to examine the processing of code-switches at open-class words (e.g., nouns) versus closed-class words (e.g., prepositions). While code-switches at both nouns and prepositions elicited a broad late positivity, only code-switches at prepositions elicited a broad early negativity, suggesting that bilinguals process code-switches differently based on their grammatical function (Zeller, 2020). Another ERP study compared how bilinguals processed code-switching at two types of open-class words: nouns and verbs (Ng et al., 2014). When reading a story, Spanish–English bilingual adults processed code-switching at nouns (e.g., "the wind and the *sol* [sp. sun]") differently than code-switching at verbs (e.g., "they *miraron* [sp. saw] a traveler") as indicated by larger N400 responses and an early Late Positive Component for nouns. The authors proposed that the difference was driven by the effort bilinguals put into integrating and remembering the information contained in each code-switch. That is, nouns are likely to be referenced several times in a story and need to be held in working memory, thus eliciting more cognitive effort compared to verbs that may only be used once. Combined, these results highlight that bilinguals may be sensitive to the functional role of the code-switched words and process them accordingly.

Research has yet to investigate how bilingual children process code-switches with diverse functional or grammatical roles, but evidence from monolinguals shows that children are sensitive to some grammatical classes beginning around 8 months of age (Marino et al., 2020).

Moreover, by age 3, children use the meaning of adjectives to predict which noun they refer to (e.g., predicting "heavy" is more likely to be followed by "stone" than "butterfly"; Tribushinina & Mak, 2016). Additionally, monolingual children as young as 2 years old can recognize, but "listen through," uninformative adjectives to quickly and correctly identify a target noun (Thorpe & Fernald, 2006). For example, when shown a picture of a dog and a bunny, children identified the target object as quickly when it was preceded by an uninformative adjective (e.g., "Where's the good bunny?") as when it was not preceded by any adjective (e.g., "Where's the bunny?"). These results show that young children can attend to the most relevant functional information to efficiently process speech.

Following the functional account, code-switching that occurs at a word that is central to the meaning of the sentence may be particularly challenging for children to process. In many cases, this will be a noun, but in other cases it could be a verb, adjective, or other part of speech. This idea is supported by previous research finding that children experience difficulty in understanding functionally-important code-switched nouns (Byers-Heinlein et al., 2017; Morini & Newman, 2019). In contrast, code-switches at parts of speech that play a limited functional role in comprehension may be relatively easy for children to process, and code-switches that are uninformative in a comprehension task may not elicit any processing difficulties. However, to date, children's comprehension of code-switches at words with limited functional meaning has not yet been investigated; thus, there is a lack of empirical evidence for the functional account with children.

4.1.3 Current Study

In the current study, we asked if code-switching within a sentence at an uninformative determiner-adjective pair (which we will hereafter refer to as an uninformative adjective) affects children's comprehension of a target noun that immediately follows it. This allowed us to compare the competing predictions of the frequency and functional accounts. The frequency account predicts that children will show disrupted processing of a code-switch at an adjective, because it is not a common location for code-switching. This could result in weaker comprehension of the following noun, as processing difficulties earlier in the sentence can negatively affect how children process the end of the same sentence (Trueswell et al., 1999). In contrast, the functional account predicts that children may find it relatively easy to process a code-switch at an uninformative adjective as they do not necessarily have to attend to or remember its meaning in the context of the visual scene.

Children viewed pairs of pictures of animals, such as a duck and a fish, and heard sentences such as "Can you find *le bon* [fr. the good] duck?" or "Can you see *el buen* [sp. the good] duck?" In trials, both animals were equally consistent with the adjective (e.g., both were depicted as equally "good"). Participants were 30 3-year-old bilinguals, including both French–English bilingual children in Montreal (*n* = 19) and Spanish–English children in New Jersey (*n* = 11). We included participants from these two testing locations to increase sample size, as bilingual children are a difficult-to-recruit population. This is in line with various sampling strategies in the field of early bilingualism which range from testing homogeneous populations (e.g., all acquiring English and French) to testing heterogeneous populations (e.g., all acquiring English and a variety of other languages; Byers-Heinlein, 2015). Assessing the effects of code-switching at adjectives was appropriate in our sample, because children of this age can generally understand their meaning (Tribushinina & Mak, 2016), and because certain adjectives can occur in the same prenominal position across the languages being acquired by our participants (i.e., English, French, and Spanish).

Similar to previous studies on children's processing of code-switching (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019), we expected that code-switching at an uninformative adjective would hinder children's comprehension of the target noun compared to sentences without code-switching. Specifically, we predicted that children would look less towards the target noun after hearing mid-sentence code-switching compared to hearing a sentence entirely in one language. Such a result would be consistent with the frequency account. In contrast, a finding that children's performance was unaffected by an uninformative code-switched adjective would support the functional account. We also explored whether individual differences such as language dominance, testing location (as a proxy for language pair), SES, or vocabulary size would be related to performance.

4.2 Methods

Data collection occurred in two locations: Montreal, Canada and New Jersey, USA. The methods were approved by the Concordia University Human Research Ethics Committee ("Monolingual and Bilingual Language Development"; approval #10000493) and the Princeton University Institutional Review Board ("Language learning and Communication"; approval

#7117), and parents provided informed consent prior to their child's participation. Data were collected in Montreal between November 2016 and April 2017 and in New Jersey between March 2017 and January 2018. Final data analysis occurred between May 2020 and June 2021, during the COVID-19 pandemic. As is common in laboratories testing hard-to-recruit populations such as bilingual children, children participated in a second, separate study, either immediately prior to or following participation in this study (the order of the two studies was counterbalanced). The results of that study are reported in a separate manuscript (Byers-Heinlein, Jardak, et al., 2021). All stimuli, data, and analysis scripts for the current study are available via the Open Science Framework at https://osf.io/ecqwr/.

4.2.1 Participants

A total of 30 3-year-old (M = 3.57, *Range* = 3.10 – 4.05, 14 females) full-term, healthy bilingual children participated in this study. This sample size was sufficiently sensitive to detect an effect size of d = 0.46 at 80% power in a paired-samples t-test, meaning there were enough participants to detect effect sizes reported in previous related studies (0.56 in Byers-Heinlein et al., 2017, 0.60 in Potter et al., 2019).

Nineteen French–English bilinguals were tested in Montreal, Canada, and 11 Spanish–English bilinguals were tested in New Jersey, USA. In Montreal, children were recruited from a database of families interested in participating in our research, principally identified via government birth lists. In New Jersey, children were primarily recruited from nonprofit organizations. Another 34 children were tested but not included in the final sample due to not meeting the language criteria (n = 15; see details below), fussiness or lack of attention (n = 10), technical issues (n = 4), health reasons such as low birth weight or gestation period under 37 weeks (n = 3), completing an insufficient number of trials (n = 1; see below), or having a reported speech delay or disorder (n = 1). Post-hoc data exclusion resulted in the unbalanced sample between the two locations. Unfortunately, because this discrepancy did not become clear until the time of data analysis, which occurred during the COVID-19 pandemic, we were unable to test additional participants to address this difference.

Children's language background and proficiency was assessed via a modified version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007; Appendix C). Parents were asked about their child's experience with the languages they were exposed to, and to rate their child's proficiency in English and French (in Montreal) or in English and Spanish (in New Jersey) compared to monolingual children of the same age. Following a pre-determined inclusion criterion, children had to receive a comprehension score of at least 7/10 for both languages to be eligible for the study. For each child, their dominant language was established as the language that had the highest comprehension score from the LEAP-Q. Twelve children had equal comprehension scores in both languages, so for these children, the language in which the child had the higher productive vocabulary score (see below) was considered their dominant language. In total, 19 children were dominant in English, 9 were dominant in French, and 2 were dominant in Spanish. Twelve children were regularly exposed to both of their languages from birth, and 18 children were exposed to their second language later in life, between the ages of 2 and 36 months. See Table 4.1 for additional information by testing location.

Children's productive vocabulary size in English was assessed using the Developmental Vocabulary Assessment for Parents (DVAP; Libertus et al., 2015; Appendix D), which consisted of a checklist of words known by children aged 2 to 18 years old based on words used in the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). We used a parent checklist rather than a direct measure to reduce children's fatigue, as each child participated in two experiments, and we wished to assess their vocabulary in both languages. Moreover, the DVAP has shown strong convergent validity with children's performance on the PPVT ($\beta = .69$; Libertus et al., 2015). To assess children's productive vocabulary size in French or Spanish, we adapted a checklist similar to the DVAP, based on words used in the adaptation of the PPVT for Quebec French (Échelle de Vocabulaire en Images Peabody; Dunn et al., 1993) or Spanish (Test de Vocabulario en Imagenes Peabody; Dunn et al., 1986). The words are ordered from easy (e.g., "ball," "dog") to hard (e.g., "honing," "angler"), and parents were asked to indicate which words their child could say. A parent or other adult that was familiar with the child's vocabulary in a particular language filled out the form for that language. In some cases, the forms for each language were completed by different parents who normally interacted with their child in that language, while in other cases it was one parent who filled out both forms if they used both languages with their child. As expected, the number of words children produced in their dominant language ($M = 71$, $SD = 32$, $Range = 24 – 177$) was greater than the number of words they produced in their non-dominant language ($M = 39$, $SD = 28$, $Range = 2 – 131$), $t(28) = 7.03$, $p < .001$, $M_d = 32.34$, 95% CI [22.92,41.77]. When combining the number of words produced in both languages, on average, children produced 110 total words (*SD* = 55, *Range* = 31 – 308). Children in Montreal (*M* = 125, *SD* = 61, range = 39 – 308) produced more words than those in New Jersey (*M* = 87, *SD* = 33, range = 31 – 138), $t(26.73) = -2.16$, $p = .040$, $\Delta M = -37.76$, 95% CI [-73.56,-1.95].

As a proxy for socioeconomic status (SES), we asked parents to indicate the highest level of education they had attained. As the education systems are somewhat different in the United States and Canada, to be able to compare responses across our two testing locations, we converted these responses to the typical number of years after kindergarten to complete each level of education (e.g., completing a bachelor's degree was equivalent to 16 years of education). For families where both parents' education was provided, the higher level was selected for analysis. On average, parents completed 15.20 (*SD* = 3.89) years of education, which ranged widely from 4 to 21 years. Parents in Montreal reported completing more years of education (*M* = 16.58, *SD* = 2.17, *Range* = 13 – 21) than parents in New Jersey (*M* = 12.82, *SD* = 5.06, *Range* = 4 – 20), $t(12.17) = 2.35$, $p = .037$, $\Delta M = 3.76$, 95% CI [0.27,7.25], suggesting that the participants in Montreal came from a higher SES background than those in New Jersey.

4.2.2 Material

4.2.2.1 Visual Stimuli

Visual stimuli consisted of 8 pairs of pictures for each language combination (see Table 4.2 for picture pairs and Figure 4.1 for an example trial). Each picture in a pair had the same animacy status (i.e., four pairs of animals used in target trials and four pairs of inanimate pictures used in filler trials), so that the two pictures had similar visual salience (see Appendix G). To ensure that they would be familiar to our 3-year-old participants, we selected pictures whose labels were highly understood by children in American English (Fenson et al., 2007), Quebec French (Boudreault et al., 2007), and Spanish (Jackson-Maldonado et al., 2003). The labels of the picture pairs did not overlap in word onset, had the same grammatical gender in French or Spanish, and are widely used across French and Spanish dialects. Pictures were chosen from free online libraries and digitally edited as necessary.

4.2.2.2 Auditory Stimuli

Auditory stimuli were recorded by a female, native French–English or Spanish–English bilingual with no perceptible accent in either language using infant-directed speech. Each auditory stimulus contained a target word labeling one of the pictures on the screen (e.g., "Look!

Can you find the good duck?"). The target noun (e.g., "duck") was preceded by a determiner (e.g., "the") and a prenominal adjective (e.g., "good"). Each stimulus sentence was recorded in a single-language version where the determiner and adjective were in the same language as the noun, and a code-switched version where the determiner and adjective were in the other language (e.g., "Look! Can you find *le bon* [fr. the good] duck?" or "Look! Can you see *el buen* [sp. the good] duck?"). Note that the target word (e.g., "duck") was always in the same language as the initial carrier phrase (e.g., "Look! Can you find…" for French–English and "Look! Can you see…" for Spanish–English). Parallel stimulus sets were created with the carrier sentences in each language (e.g., in French, the previous examples became "*Regarde! Peux-tu trouver le bon canard*?" and "*Regarde! Peux-tu trouver* the good *canard*?"; in Spanish, the previous examples became "*¡Mira! Puedes ver el buen pato*?" and "*¡Mira! Puedes ver* the good *pato*?").

For the animate nouns on target trials, there were a total of four English prenominal adjectives and their French and Spanish translations; similarly, for inanimate nouns in filler trials, there were four prenominal adjectives used (see Table 4.2). These adjectives were chosen such that they 1) were not cognates across French and English or Spanish and English, 2) did not share phonological overlap with their translation, 3) were not descriptive of one picture more than another, and 4) could precede a noun in French or Spanish. Although both French and Spanish usually place adjectives in a postnominal position, the adjectives we selected can be used prenominally in these grammatical contexts. Each adjective was always used with the same picture pair.

4.2.2.3 Trial Description

During each trial, the target and distractor pictures appeared on the screen for 6000 ms, and one of the stimulus sentences was played labeling the target picture. The onset of the target noun occurred exactly 3000 ms into each trial. The determiner–adjective pairs were of somewhat different lengths, and so occurred between 311 and 1152 ms before the noun onset. Trials were combined into four experimental orders of 24 trials: 8 single-language trials (e.g., "Look! Can you find the good duck?"), 8 code-switched trials (e.g., "Look! Can you find *le bon* [fr. the good] duck?"), and 8 additional single-language filler trials. Filler trials were not analyzed and were mainly used to lower the overall number of trials with code-switching. Target trials (i.e., single-language and code-switched trials) and filler trials were intermixed throughout the study. The language of the carrier phrase was consistent for each child (i.e., always in English, French, or

Spanish), but counterbalanced across children at the time of testing. In total, 15 children were tested with carrier phrases in their dominant language (10 French–English and 5 Spanish–English), and 15 children were tested with carrier phrases in their non-dominant language (9 French–English and 6 Spanish–English).

4.2.3 Procedure

In addition to signing a consent form (Appendix B), parents completed questionnaires on their child's language comprehension (LEAP-Q; Appendix C) and vocabulary (DVAP; Appendix D), on their own language mixing (Byers-Heinlein, 2013; Appendix E), and on basic demographic information (Appendix F). During the study, parents listened to music with headphones, wore darkened glasses, and were instructed not to interfere with the study or provide their child with any instruction. Testing occurred in a darkened room while children sat on their parent's lap.

Due to differences in lab equipment, the same apparatus was not available at both testing sites. In Montreal, the study was conducted in the lab on a 24-inch Tobii T60XL corneal reflection eye-tracking system using a 5-point calibration, with auditory stimuli played over speakers. In New Jersey, the study was conducted either in the lab (7 children) or at a local community center (4 children), depending on which location was easier for participants to access. In the lab, the study was run on a 55" TV monitor while the auditory stimuli were played over speakers. At the community center, children completed the study on a 13" laptop while listening to the stimuli over noise-canceling headphones. In both New Jersey setups, a video camera below the screen recorded children's eye movements at a rate of 30 frames per second for later offline coding by trained research assistants.

Before each trial began, a colorful attention-getter was presented to draw the child's attention to the screen. Once the child was looking at the screen, the trial began. An experimenter monitored the status of the study via video camera and controlled the experiment from a computer in another room (Montreal) or within the same room (New Jersey). The total duration of the study was approximately 4 minutes.

4.2.4 Coding

In Montreal, the eye-tracking system collected data on the location of children's eye-gaze and their pupil size at a rate of 60Hz. We defined areas of interest corresponding to a rectangle of 2 cm around each picture presented on the screen. In New Jersey, a trained research assistant manually coded videos with frames at 33-ms intervals for whether the child was looking at the left or right object on the screen, shifting between objects, or inattentive. A second research assistant coded 18% of videos; on the frames surrounding eye movements, inter-coder reliability was 97%. Research suggests that automatic eye-tracking and manual gaze coding, although potentially different in their amount of data loss, capture largely similar information (Venker et al., 2020). We did not observe a difference in data loss between the two coding methods. An average of 15.88% (SD = 9.31) of eye-tracking data and 15.59% (SD = 8.16) of manually coded data was lost for each participant, $t(23.37) = 0.09$, $p = .929$, $\Delta M = 0.00$, 95% CI [-0.06,0.07]. Additionally, previous research has combined data across these methods to create a single bilingual sample (Byers-Heinlein, Jardak, et al., 2021), further supporting this approach.

4.3 Results

Data for each trial were analyzed between 400 and 2000 ms after the onset of the target noun. While standard approaches typically begin analysis at 367 ms after onset of the target noun (Swingley, 2012), we opted to start our analysis window slightly later in order to create consistent 100-ms time bins to use in a growth curve analysis (see below). Trials where the child was inattentive (i.e., looked at the pictures for less than 750 ms during this window) were excluded from the analyses. Children who did not successfully complete at least 2 single-language and 2 code-switched trials were also removed from the analyses. Out of 8 possible trials of each type, children retained for analysis completed an average of 6.87 single-language trials (*Range* = 3 – 8) and 6.63 code-switched trials (*Range* = 4 – 8). To determine if children demonstrated successful comprehension of the target words, we examined the proportion of time that they looked towards the target picture on each trial. This was calculated by dividing the looking time to the target picture by the total time spent looking at either picture.

First, we investigated whether children showed comprehension of the noun on each trial type. One-sample, two-sided t-tests revealed that children looked significantly above chance ($\mu_0 = 0.5$) to the target picture on both single-language trials, $t(29) = 11.42$, $p < .001$, $M = 0.74$, 95% CI [0.70,0.78], and code-switched trials, $t(29) = 12.03$, $p < .001$, $M = 0.78$, 95% CI [0.73,0.82], indicating a robust ability to understand the target noun in both trial types (see Figure 4.2).

We then compared looking time during the two trial types using a paired-samples t-test. The effect of trial type was not statistically significant, $t(29) = 1.49$, $p = .148$, $M_d = 0.04$, 95% CI [-0.01,0.09], suggesting that children's comprehension of the noun did not differ between single-language and code-switched trials. Contrary to our prediction that children's comprehension of the target noun would be impaired by the code-switching that preceded it, this result indicated that they were potentially unaffected by the code-switched adjective.

4.3.1 Growth Curve Analysis

The previous analyses, which are typical in this area of research, collapsed infants' data across the entire time window and averaged across trial types to yield two data points per child. However, it has long been recognized in the field that time course data can offer revealing information about children's performance (e.g., Fernald, Swingley, et al., 2001). Analytic techniques such as growth curve analysis (Mirman, 2017) offer an approach to quantify differences in time course, and further allow analysis of trial-level data, thus increasing statistical power. We plotted the time course of our data and then conducted an exploratory growth curve analysis, using the same time window of 400 – 2000 ms. Looking-time data were binned in 100-ms blocks.

Models were built through an iterative process. We started with a baseline model with only linear and quadratic time terms and by-participant random effects on both time terms. We then added one additional individual difference variable to the model and compared the two nested models with an analysis of variance. Only variables that significantly improved model fit were retained. Intermediary models are available in Appendix H. The categorical variables of trial type, testing location, and language dominance were coded using a simple contrast coding scheme. SES and vocabulary size were continuous. We estimated parameter estimate degrees of freedom and *p*-values using Satterthwaite's method.

To address our main research question of the effect of code-switching on children's comprehension, our first exploratory model added trial type to the baseline model described above. We then conducted additional exploratory growth curve models building from this model looking at the potential individual effects of language dominance, testing location, SES, and vocabulary size.

4.3.1.1 Trial type

In the growth curve model investigating the effect of trial type, the fixed effects of the final model included trial type, and linear and quadratic time terms. There was a statistically significant main effect of trial type, indicating that, opposite to our prediction, children were more accurate at gazing toward the target picture when hearing code-switched trials compared to single-language trials, $t(6{,}100.82) = -3.43$, $p = .001$, $\hat{\beta} = -0.03$, 95% CI [-0.05,-0.01] (see Table 4.3 for full results). This result differs from that of the paired-samples *t*-test, which did not find a statistically significant difference in children's looking between the two trial types.

4.3.1.2 Individual Differences

As previous studies have found some evidence of individual differences in bilingual children's ability to process code-switching (Byers-Heinlein, Jardak, et al., 2021; Potter et al., 2019), we next investigated how such differences may have affected children's performance on this task. Prior to conducting these individual differences analyses, we first quantified the consistency of children's performance, by estimating the reliability of the looking time to each trial type using an intraclass correlation coefficient (ICC), based on a mean-rating, consistent, 2-way random-effects model (Byers-Heinlein, Bergmann, et al., 2021). The estimated consistency was 0.19, 95% CI = [–0.24, 0.51] for single-language trials and 0.39, 95% CI = [0.07, 0.64] for code-switched trials. The magnitude of these ICCs was higher than in many other infant studies (Byers-Heinlein, Bergmann, et al., 2021), supporting a cautious investigation of individual differences. However, these ICCs could be considered moderate to low on an absolute scale thus reducing statistical power for detecting correlations with other measures of individual differences.

We investigated four individual difference variables: language dominance, testing location (which was also a proxy for language pair), SES, and vocabulary size. We note that the last three variables were interrelated in our dataset: children from Montreal generally came from higher SES backgrounds, $t(12.17) = 2.35$, $p = .037$, $\Delta M = 3.76$, 95% CI [0.27,7.25], and had a larger vocabulary, $t(26.73) = -2.16$, $p = .040$, $\Delta M = -37.76$, 95% CI [-73.56,-1.95], than children from New Jersey. Given our sample size, it was not possible to statistically disentangle these factors. Thus, our approach was to create separate models for each variable to gain some insight into which factor might have the largest explanatory power. We did so by adding each variable to the previous model including trial type as a main effect and in an interaction with trial type. Here, we focus on the specific effect of these terms. Full results of these models are reported in Appendix H.

In each model, there was a statistically significant main effect of trial type, indicating that, opposite to our prediction, children were more accurate at gazing towards the target picture when hearing code-switched trials compared to single-language trials, whether controlling for language dominance, $t(6,101.58) = -3.39$, $p = .001$, $\hat{\beta} = -0.03$, 95% CI [-0.05,-0.01], testing location, $t(6,103.15) = -4.67$, $p < .001$, $\hat{\beta} = -0.05$, 95% CI [-0.07,-0.03], SES, $t(6,106.01) = -4.75$, $p < .001$, $\hat{\beta} = -0.20$, 95% CI [-0.28,-0.12], or vocabulary, $t(5,899.58) = -2.10$, $p = .035$, $\hat{\beta} = -0.05$, 95% CI [-0.10,0.00].

We then examined the main effect of each individual difference variable and its interaction with trial type (see Figure 4.3), and an interesting pattern of results emerged. For language dominance, there was no statistically significant main effect, $t(29.44) = -1.36$, $p = .183$, $\hat{\beta} = -0.05$, 95% CI [-0.11,0.02], or interaction with trial type, $t(6,101.58) = 0.35$, $p = .727$, $\hat{\beta} = 0.01$, 95% CI [-0.03,0.05], suggesting that children tested in their dominant language and children tested in their non-dominant language performed similarly across trial types. Effects of testing location, SES, and vocabulary showed similar patterns across models. Analyses of testing location revealed that children from Montreal performed similarly on both trial types, whereas children from New Jersey performed better on code-switched than single-language trials $t(6,103.14) = -4.16$, $p < .001$, $\hat{\beta} = -0.09$, 95% CI [-0.13,-0.05]. To follow up on the Montreal results, we conducted the pupillometry analyses reported in Appendix H, which support the main finding that children did not process code-switched and single-language trials differently (these analyses could not be carried out for New Jersey participants, as their data were hand coded from a video recording rather than collected via an eye-tracker). SES analyses showed that children from higher-SES backgrounds performed similarly across trial types whereas children from lower-SES backgrounds performed better on code-switched than single-language trials, $t(6,103.72) = 4.04$, $p < .001$, $\hat{\beta} = 0.01$, 95% CI [0.01,0.02]. Finally, children with larger vocabularies performed better across trial types (i.e., looked more to the labeled target in general) than children with smaller vocabularies, $t(28.38) = 2.42$, $p = .022$, $\hat{\beta} = 0.0007$, 95% CI [0.0001,0.0013], but the effect of vocabulary size did not differ significantly as a function of trial type, $t(5,896.30) = 0.85$, $p = .396$, $\hat{\beta} = 0.0002$, 95% CI [-0.0002,0.0005].

These results indicate that individual differences in performance across the two trial types were statistically related to testing location and SES, but not to language dominance or vocabulary size. Spanish-English bilingual children from New Jersey, particularly those whose parents had received a high school education or less (i.e., 12 years or fewer; see Figure 4.3), performed better on code-switched trials compared to single-language trials, whereas French-

English bilingual children and those whose parents had more educated performed similarly on the two trial types. Together, the findings show the importance of examining individual differences between participants and samples, as bilingual children's comprehension of these code-switched sentences was not uniform.

4.4 Discussion

This study compared bilingual children's comprehension of sentences with code-switching at an uninformative determiner-adjective pair (e.g., "Can you find *le bon* [fr. the good] duck?") to their comprehension of single-language sentences (e.g., "Can you find the good duck?"). We tested 3-year-old bilingual children, including French–English bilinguals in Montreal and Spanish–English bilinguals in New Jersey. We found that bilinguals were, on average, successful at identifying the target noun in both types of sentences, and we did not see evidence that code-switching at an uninformative adjective caused any difficulties in sentence processing. Language dominance did not affect performance, likely because the target noun was always presented in a consistent language, and the switch occurred at the preceding adjective. This finding contrasts with prior reports of dominance effects in studies of children's processing of code-switches (Potter et al., 2019). Surprisingly, we found some evidence that, for certain children, code-switched sentences may have facilitated comprehension relative to single-language sentences.

Our experimental design allowed us to test two sets of competing theoretical predictions. Under the frequency account of code-switch processing, the infrequent nature of code-switching at determiner-adjective pairs should have hindered children's comprehension, perhaps even more so than code-switching at nouns (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019). In contrast, under the functional account, children may have been able to seamlessly process code-switching at an uninformative adjective, because they did not need to integrate the meaning of the adjective to identify the target noun. Our results generally support the functional account as children were able to understand the code-switch sentences as well as the single-language sentences. Below, we further discuss why young children's processing was not disrupted by code-switching at uninformative adjectives. Then, we turn to addressing the observed individual differences between participants and communities.

A key aspect of our experimental design was that the determiner-adjective pair in our sentences was uninformative. Children heard sentences with mid-sentence code-switching, as in

"Can you find *le bon* [fr. the good] duck?" Critically, the adjective "*bon*" [fr. good] did not add relevant information for identifying the target object, as there was only one duck on the screen. Children typically process the meaning of adjective–noun phrases incrementally (Fernald et al., 2010; Tribushinina & Mak, 2016), but they can "listen through" the adjective to quickly identify the target object when a prenominal adjective is uninformative and does not disambiguate two objects (Thorpe & Fernald, 2006). Following the functional account, code-switching may not be disruptive when the information it carries does not need to be retrieved or integrated into processing. Children may not have experienced a code-switching cost in the current study because they did not need to process the meaning of the code-switched adjective to identify the target and were therefore able to ignore it.

Similarly, if code-switching is related to prediction processes during language comprehension (e.g., Yacovone et al., 2021), the unexpected code-switch at the adjective might have led to a brief processing slowdown combined with a simultaneous increase in attention (Reuter et al., 2019), effectively canceling each other out in the context of an uninformative adjective. Thus, derailment in children's processing of code-switches may be limited to functionally important words or phrases that require them to integrate the information contained in the switch.

To further test this possibility, future studies could compare performance on trials like those in the current study and trials with an informative adjective (e.g., by showing a picture of a big and small duck and examining children's real-time interpretation of the sentence "Do you see *le petit* [fr. the little] duck?"). Under the functional account, sentences with an informative adjective would presumably result in a code-switching cost, because children would no longer be able to "listen through" the code-switched adjective and would potentially need to engage their other language more fully.

While "listening through" could explain why we did not observe a code-switching cost in this study, it does not explain the observed individual differences in children's performance on code-switched and single-language sentences. Our analyses revealed that testing location and SES accounted for significant individual variation in performance across the single-language and code-switched trials, but language dominance and vocabulary size did not. Specifically, children from higher-SES backgrounds performed similarly across trial types; children from lower-SES backgrounds, particularly whose parents received a high school education or less, performed better on code-switched trials than single-language trials, and were all Spanish–English bilinguals in New Jersey.

In our sample, testing location (a proxy for language pair), SES, and vocabulary size were tightly related: French–English children from Montreal had higher vocabularies and were from higher SES backgrounds on average than Spanish–English children from New Jersey. Because of the correlational nature of this finding and the interrelatedness of these variables, it is not possible to pinpoint the factors driving the individual differences we observed. However, previous studies have reported similar patterns of individual differences in infants from these same communities; one study suggested that Spanish–English children may have slightly weaker skills in real-time language tasks than French–English children (Byers-Heinlein, Jardak, et al., 2021). Following the functional account, if some children were slower to switch between processing their two languages, or if they were less aware of its meaning, it is possible that they were able to "listen through" the uninformative adjective more easily (or under a prediction-based framework, encountered little to no prediction error). However, note that under this explanation, we would have expected vocabulary size to predict performance, which it did not. Rather, SES was a predictor of performance, a variable which has previously been related to children's language development (Fernald et al., 2013; Pace et al., 2017; Pungello et al., 2009). We tentatively suggest that experiential factors related to SES might be driving the observed community differences.

There are also other potentially relevant differences between children that we were not able to directly observe that may have affected infants' performance on our task. For example, different infants have different experiences with code-switching (Bail et al., 2015; Chapter 3), which could in turn impact their comprehension of code-switching. The frequency account predicts that bilinguals with frequent exposure to code-switching should experience less disruption in processing compared to bilinguals without frequent exposure to code-switching (Gosselin & Sabourin, 2021; Valdés Kroff et al., 2018). In the context of the current study, experience with code-switching may have been able to build on top of children's ability to "listen through" the uninformative adjective. It is also possible that production of code-switching varies by SES within the two communities we studied, although this has not yet been examined directly. We speculate that Spanish-English bilinguals in New Jersey, particularly those from lower-SES backgrounds, may have been more accustomed to hearing code-switching than our other participants, resulting in the potential boost in real-time sentence interpretation – at least in the context of sentences with mid-sentence code-switches at uninformative locations. To address this question, additional research is needed to directly investigate the relationship between the amount and type of code-switching that bilingual children hear and how they process incoming speech input in two languages.

4.5 Conclusion

Code-switching is common in bilingual speech, making it important to understand its effect on children's language comprehension and language learning. Past research has generally found that code-switching leads to processing costs, but in the current study, bilingual children did not show this processing cost. They showed similar (and in some cases, better) processing of sentences with a code-switch at an uninformative adjective phrase, relative to single-language sentences. These findings demonstrate that linguistic features such as informativeness and location may impact how bilingual children process code-switching in natural settings.

Table 4.1: Demographics of participants at each testing location.

Testing Location	Total *n*	Mean age in years (*Range*)	English dominant (*n*)	L2 exposure from birth (*n*)	Later L2 exposure (age range in months)	Dominant Language Vocabulary (*SD*)	Non-Dominant Language Vocabulary (*SD*)	Parental education (*SD*)
Montreal	19	3.47 (3.1 – 3.99)	10	8	6 – 18	76.83 (33.91)	47.83 (30.19)	16.58 (2.17)
New Jersey	11	3.75 (3.19 – 4.05)	9	4	2 – 36	62.36 (26.22)	24.55 (18.34)	12.82 (5.06)

Note. English dominant (*n*) lists the number of children at each testing location who were dominant in English; the remainder of children were dominant in either French if tested in Montreal or Spanish if tested in New Jersey. Later L2 exposure (age range in months) only considers participants who were not exposed to both languages from birth.

Table 4.2: Adjective–noun pairs used for French–English and Spanish–English participants.

English		French	
Look! Can you find … ?		*Regarde! Peux-tu trouver … ?*	
Adjective	**Noun pair**	**Adjective**	**Noun pair**
Target trials			
the good	duck – fish	le bon	canard – poisson
the little	monkey – sheep	le petit	singe – mouton
the nice	dog – bunny	le gentil	chien – lapin
the pretty	cow – froggy	la jolie	vache – grenouille
Filler trials			
a large	ear – spoon	une grosse	oreille – cuillère
a new	apple – toothbrush	une nouvelle	pomme – brosse à dents
a big	door – hand	une grande	porte – main
an old	coat – pencil	un ancien	manteau – crayon

English		Spanish	
Look! Can you see … ?		*¡Mira! ¿Puedes ver … ?*	
Adjective	**Noun pair**	**Adjective**	**Noun pair**
Target trials			
the good	bear – duck	el buen	oso – pato
the little	butterfly – sheep	la pequeña	mariposa – oveja
the big	bunny – dog	el gran	conejo – perro
the pretty	cow – froggy	la hermosa	vaca – rana
Filler trials			
a beautiful	ear – spoon	una linda	oreja – cuchara
a new	apple – toothbrush	una nueva	manzana – cepillo de dientes
a nice	door – hand	una preciosa	puerta – mano
an old	coat – pencil	un viejo	chamarra – lápiz

Note. The noun pairs labelled the two pictures shown on screen at the same time. Each noun was used as a target in different trials. In single-language trials, the adjective and noun were in the same language. In code-switched trials, the adjective and the noun were in different languages.

Table 4.3: Growth curve analysis including trial type.

	Estimate	95% CI	*t*	*df*	*p*
Fixed effects					
Intercept	0.76	[0.72, 0.79]	43.05	29.42	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.86	29.46	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.36	29.09	< .001
Trial type	-0.03	[-0.05, -0.01]	-3.43	6,100.82	.001
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.154			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + ([time (linear) + time (quadratic)] || participant)*

Figure 4.1: Example and timeline of experimental trial.

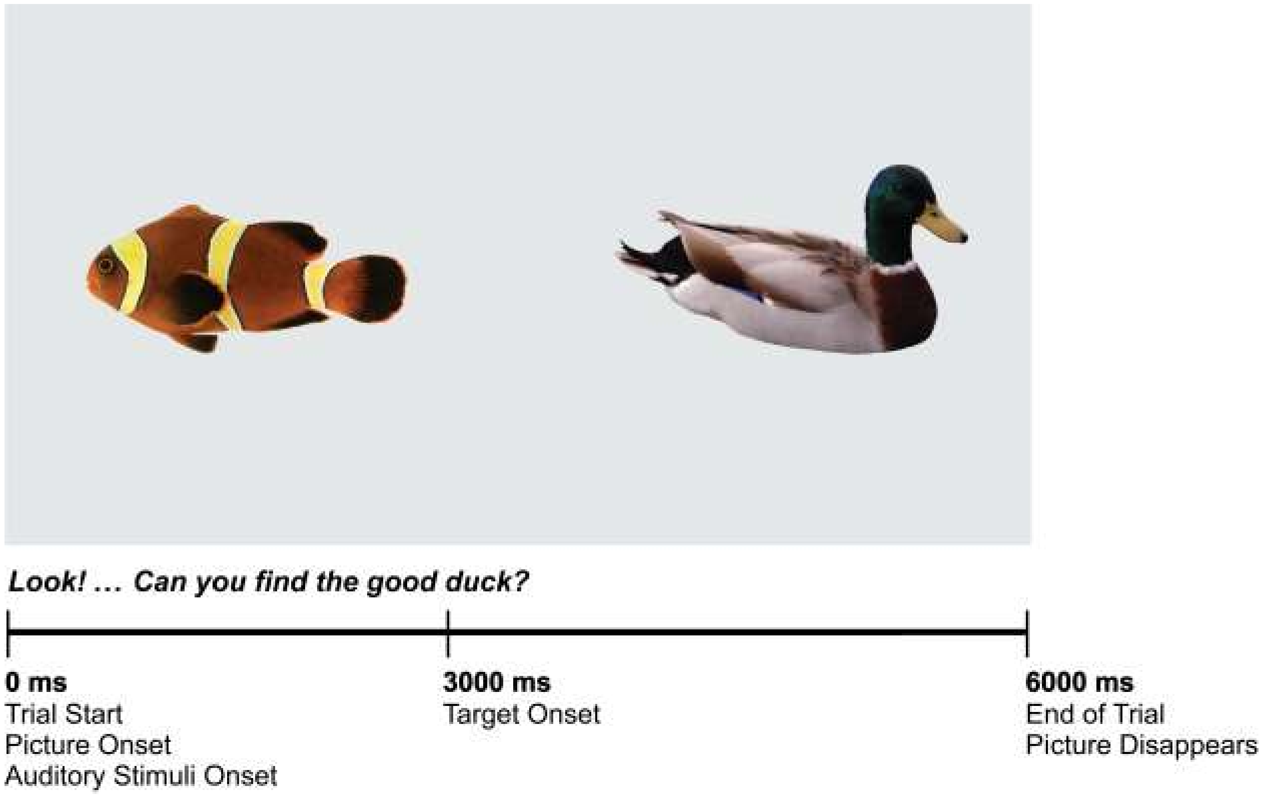
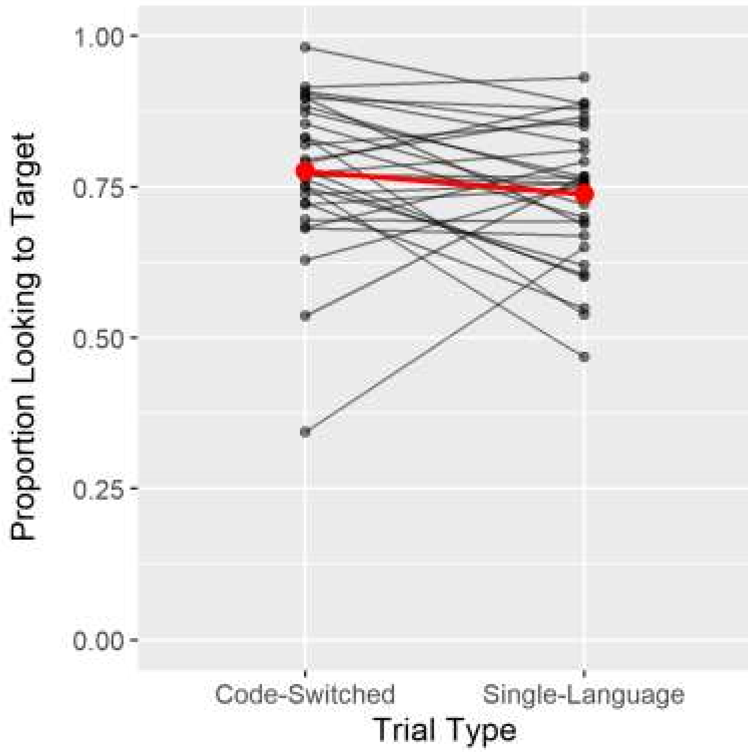
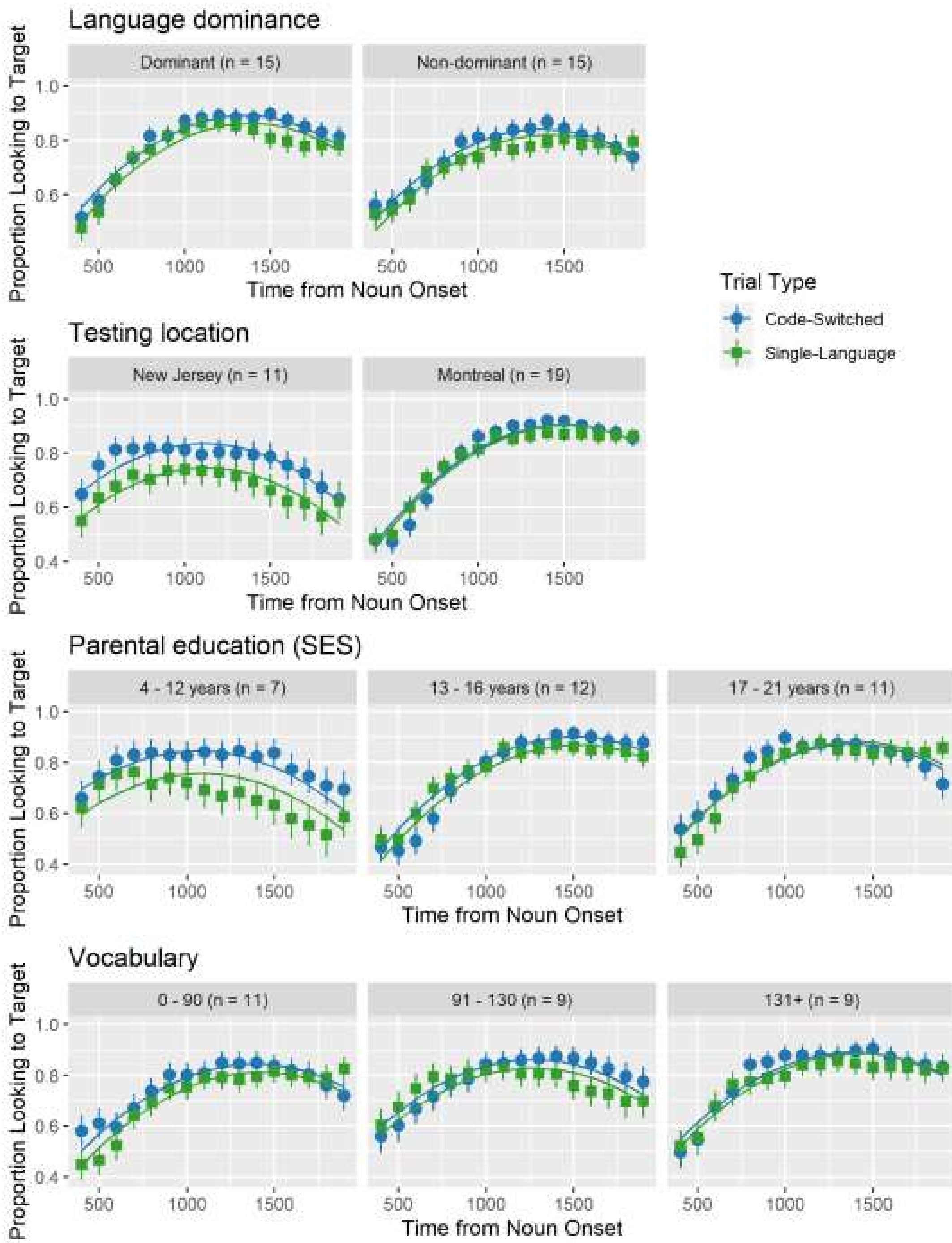


Figure 4.2: Proportion looking to target picture by trial type for all children.



Note. The larger red dots and line represent the grand mean. Smaller gray dots and their connecting lines represent the mean values for individual participants.

Figure 4.3: Proportion looking to target picture throughout the analysis window.



Note. Dots represent means averaged over participants, bars represent ± 1 SEM, and lines represent the growth curve analysis model. SES and vocabulary were included in the model as a continuous variable but have been split into categories for the purposes of visualization. Note that one participant did not have a vocabulary score and was thus excluded from that model.

5 General Discussion

This dissertation had two main objectives. The first objective was to evaluate current approaches to defining and modeling bilingualism. The second objective was to examine the code-switching in bilingual children's environment and its impact on their language comprehension. Both objectives fit within the overall theme of the dissertation of better understanding bilingualism, particularly within a developmental context.

Addressing the first objective, Chapter 2 reviewed the two main current practices used in the field which model bilingualism as either a categorical or continuous variable. I identified two psychometric models that could be used to integrate both categorical and continuous information into a single model of bilingualism: the factor mixture model and the grade-of-membership model. Both models offer the nuance of a continuous model while simultaneously allowing researchers to identify separate groups of bilinguals. Using these models allows researchers to analyze their data using a categorical, continuous, or combined approach depending on the specifications of the model and their research question. The widespread use of these models could allow findings to be more easily synthesized across studies. Nonetheless, Chapter 2 cautions against a single model of bilingualism being expected across the entire field, as different models may be more appropriate for different populations or research questions.

Addressing the second objective, Chapters 3 and 4 focused on better understanding code-switching in bilingual children's environment and how it may affect their language processing and comprehension. Chapter 3 evaluated the code-switching that bilingual infants hear in their daily lives from their parents. Analyzing daylong home audio recordings from French–English families in Montreal revealed that the frequency of code-switching varied between families and generally increased across the infant's development. The majority of the code-switches that parents produced occurred intersententially (e.g., "Come on. *C'est fini*. [fr. It's done.]"), as opposed to intrasententially (e.g., "*Est-ce qu'on va aller manger* [fr. Are we going to eat] banana pancake?"). Additionally, parents appeared to code-switch most often to bolster their infant's understanding and to teach more new words as their infant's vocabulary grew. Chapter 4 investigated how 3-year-old bilingual children process and understand code-switching across two testing sites. Bilingual children did not display any difficulties in processing or understanding sentences with a code-switch at an uninformative determiner-adjective pair (e.g., "Can you find *le bon* [fr. the good] duck?) compared to single-language sentences without a code-switch (e.g., "Can you find the good duck?"). Surprisingly, some children appeared to process code-switched sentences better than single-language sentences. Exploratory analyses suggested that individual differences, such as SES and language pair, could explain these unexpected results.

Below I examine the main contributions that this work offers to the field of bilingualism, focusing on the role of code-switching in children's language development and the methodological advancements made in this dissertation. I further explore the broader implications of the main findings, including the importance of incorporating naturally produced code-switching and individual differences in study designs and the consequences of code-switching on young bilinguals' language separation. Finally, I discuss limitations and ideas for future research that build off the findings of this dissertation.

5.1 Main Contributions

This dissertation makes two main sets of contributions to the fields of bilingualism and language development. The first set of contributions regards an increased understanding of the role of code-switching in bilinguals' language development. The findings from Chapter 3 suggest that parents' code-switching may adapt to and support their child's bilingual language development; the findings from Chapter 4 suggest that how children process code-switching may be related to the functional role of the code-switch and the frequency children hear code-switching. The second set of contributions is in the form of two methodological advances that allow bilingualism and language development to be studied in a richer and more nuanced way. Chapter 3 developed a new protocol to analyze naturally produced code-switching, and Chapter 2 proposed the use of psychometric models that allow researchers to simultaneously analyze bilingual data categorically and continuously.

5.1.1 Role of Code-Switching

This dissertation investigated bilingual language development through the lens of code-switching, a distinctive property of bilingual speech. Chapter 3 was uniquely able to look at how parents' naturally produced code-switching changes across their infant's development. Parents' code-switching appeared to change in response to their infant's development, particularly around the milestone of the infant's first words. Bilingual parents altering their code-switching is in line with previous findings that monolingual parents modify properties of their speech, such as pitch (Kitamura & Burnham, 2003; Kitamura & Lam, 2009; Stern et al., 1983), vowel articulation (Lam & Kitamura, 2012), and syntactic complexity (Elmlinger et al., 2019), to support their child's language development. Moreover, this finding demonstrates that parents' code-switching is a dynamic property of their speech and should thus be studied developmentally. One way that parents' code-switching may change as their child's language skills improve is the frequency of their code-switching. I found that parents code-switched more, particularly intersententially, when their infant was 18 months old than when they were 10 months old. Another way parents' code-switching seems to change is through the apparent reasons that motivate their code-switching. Previous research has reported many different reasons why parents code-switch but did not investigate how often parents' code-switching is motivated by these different reasons (Byers-Heinlein, 2013; Goodz, 1989). I found that parents code-switched most frequently to support their child's understanding and that they code-switched more to teach vocabulary after the age at which children typically say their first words. These results suggest that as children's language skills develop, code-switching could be another feature of parents' speech that they adapt to provide more learning opportunities and support their child's bilingual language development.

Chapter 4 looked at how young bilinguals process the code-switching that they hear and was the first study to investigate children's processing of code-switches that occurred at a syntactic location other than the target noun. By presenting a code-switch at a prenominal, uninformative determiner-adjective pair, I was able to evaluate two accounts of code-switching processing: the frequency account (Gosselin & Sabourin, 2021; Valdés Kroff et al., 2018; Vaughan-Evans et al., 2020) and the functional account (Ng et al., 2014; Zeller, 2020). The results generally support the functional account, which hypothesizes that the role that a code-switch plays and the information it carries affects how it is processed. Children were not affected by the code-switching and appeared to be able to "listen through" the uninformative adjective (Thorpe & Fernald, 2006). However, the results from exploratory analyses also appear to support the frequency account, which hypothesizes that bilinguals who hear code-switching frequently in their input are better at processing code-switching than those who hear code-switching infrequently. Spanish–English children, who may have heard more code-switching than the

French–English children, performed better when hearing code-switched than single-language trials. Overall, these findings suggest that not all code-switching is difficult for children to understand, contrary to previous hypotheses (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019), and that children's comprehension of code-switching could be influenced by a combination of properties of the code-switch itself and individual differences between children.

5.1.2 Methodological Advancements

This dissertation also proposed two new methodologies to investigate and incorporate more variation in bilinguals' experiences into study designs, which will ultimately enrich the understanding of bilingualism. First, Chapter 3 provides a new protocol to analyze naturally produced code-switching to understand both its frequency and properties. This method, which relies on audio recordings, can reveal the code-switching that young bilinguals hear in their daily lives. Additionally, this method has several advantages relative to previous methods, which have included questionnaires (Byers-Heinlein, 2013), diaries (Place & Hoff, 2016), and laboratory studies (Bail et al., 2015). Most centrally, the method I developed does not rely on parents' accurate self-assessment, and thus it reduces the chances that parents may alter their code-switching based on their perceived expectations of the environment (Dewaele, 2010; Dewaele & Li, 2014; Ritchie & Bhatia, 2012). This is an important advancement in the field of bilingual language development, because it allows researchers to gain a deeper understanding of the language input that young bilinguals hear in their daily lives. As I argued in Chapter 1, children's language input is a crucial component in their language development. Thus, the protocol developed for Chapter 3 can help to better understand young bilinguals' input and ultimately advance the understanding of the process of bilingual language development.

Second, Chapter 2 is the first to suggest that bilingualism researchers move away from relying solely on categorical or continuous models of bilingualism and instead approach modeling bilingualism through a combined approach. This approach has provided new insights into various fields, including psychological disorders (Clark et al., 2013; Hallquist & Wright, 2014), voting patterns across political parties (Gormley & Murphy, 2009), and sport psychology (Brown et al., 2017). If applied to research on bilingualism, using either the factor mixture model or the grade-of-membership model could greatly deepen our understanding across the spectrum of bilinguals' experiences. Similar to calls for standardization in the field (e.g., De Cat et al., 2021; Marian & Hayakawa, 2021), if these models are adopted widely in the field, and the same model is used across multiple studies, it will allow for easy comparison and consolidation of results. The adoption of these models would lead to a richer understanding of bilingualism and the many different forms that it can take.

5.2 Broader Implications

Combined, this work has at least three broader implications for the field of bilingual language development. First, it highlights the need for additional work examining the role of naturally produced code-switching in bilinguals' language development. Second, it emphasizes how researchers should account for more individual differences to better capture the diversity of bilinguals' experiences and to better understand how these differences impact language development. Lastly, it points to how code-switching offers a unique lens through which to study language development. In particular, I will discuss how studying code-switching reveals the process of language separation and how it might play a unique role in this process.

5.2.1 Incorporating Naturally Produced Code-Switching

Code-switching is a part of bilingual children's daily lives, and the work presented in this dissertation reveals the importance of studying the impact of naturally produced code-switching on bilingual language development. This work is in line with broader trends in the field of language development, where researchers are increasingly using audio recording methods to collect naturalistic language input data from larger samples of children (e.g., Räsänen et al., 2021; Soderstrom et al., 2021; Warlaumont et al., 2017), building on a rich history of case-studies (MacWhinney, 2014; Yip et al., 2018). The coding protocol developed in Chapter 3 could help researchers identify patterns in the code-switching that young bilinguals hear in their language input, particularly in different bilingual populations. Once it is understood what code-switching young bilinguals hear, researchers can then turn to focusing on how this code-switching affects their language development. Two particularly relevant elements of language development to investigate in relation to naturally produced code-switching are vocabulary development (Bail et al., 2015; Byers-Heinlein, 2013; Carbajal & Peperkamp, 2020; Place & Hoff, 2016) and language comprehension (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019).

The findings from Chapter 3 suggest that parents' naturally produced code-switching may support their child's vocabulary development. This hypothesis contrasts with previous results, which found either a negative or no relationship between parents' code-switching and their child's vocabulary size, and likely stem from the different methods used to measure parents' code-switching (Bail et al., 2015; Byers-Heinlein, 2013). Chapter 3 measured the frequency and apparent reasons behind parents' code-switching in their natural, daily home environment, eliminating parents' self-evaluation or potential changes in their behavior when under direct observation in the lab. Bilinguals adapt their use of code-switching to their environment (Dewaele, 2010; Dewaele & Li, 2014; Ritchie & Bhatia, 2012), and parents may change the way they interact with their child in a laboratory setting (Bornstein et al., 2006; Leyendecker et al., 1997; Stevenson et al., 1986). Thus, at-home recordings likely provide a more accurate assessment of young bilinguals' daily language environment than laboratory-based observation. Using at-home recordings in Chapter 3 also provided the opportunity to determine why a parent may have produced a particular code-switch. Because parents were found to code-switch frequently to support their child's understanding and to teach new words, parents' code-switching may bolster their child's vocabulary in each of their languages. This hypothesis was not directly evaluated in this dissertation, so future research comparing parents' naturally produced code-switching and the child's vocabulary size is needed to examine these relationships. Such research would offer a unique perspective on bilinguals' language development, as it would be evaluating the effect of the daily language input that young bilinguals receive.

The results of Chapter 4 suggest that not all code-switching is difficult for children to understand, which is consistent with recent findings that bilingual adults do not experience processing difficulties when code-switching is evaluated in more naturalistic contexts (Blanco-Elorrieta & Pylkkänen, 2018). In laboratory studies, conditions are understandably highly controlled to isolate the variable of interest. This level of control can lead researchers to present stimuli that may not be representative of the code-switching that bilinguals encounter in their daily lives. The stimuli may contain code-switches with unnatural syntactic structures or a lack of contextual cues. Thus, toddlers' code-switching difficulties seen in previous laboratory studies (Byers-Heinlein et al., 2017; Morini & Newman, 2019; Potter et al., 2019) may not be present when children process code-switching in their daily lives. Although the code-switch presented in Chapter 4 may have been infrequent in the bilinguals' input, the results still suggest that presenting code-switches at different syntactic locations can lead to different results. Future research could first determine what types of code-switching children hear in their daily lives and then compare how children process code-switches that are frequent versus infrequent in their environment.

By first investigating the code-switching that young bilinguals hear in their daily lives and then using this information to examine how they process different types of code-switches, researchers will gain a better understanding of the role that code-switching plays in bilinguals' language development. This line of work would also benefit from accounting for variation between bilinguals' experiences with code-switching, as these individual differences may influence language development.

5.2.2 Incorporating Individual Differences

The idea that bilingualism is not a monolithic experience is not new (e.g., de Bruin, 2019), and each of the chapters in this dissertation have discussed individual differences in bilinguals' experiences and abilities. Chapter 2 described why it is important to account for individual differences and proposed the integration of new models that allow for such variation in approaches to bilingualism. Chapter 3 found that families code-switched at different rates and for different reasons. Chapter 4 revealed that children's comprehension of code-switching varied across factors such as testing location and SES. These findings further confirm the importance of accounting for and investigating the effects of individual variation on bilingual language development. To date, most research on bilingual language development has been limited in the extent to which it is able to investigate how individual differences may affect bilinguals' language development, because researchers have generally relied on a categorical approach to bilingualism where all bilinguals in a given study are analyzed as a single group. Moving away from a purely categorial view of bilingualism requires researchers to consider what individual differences they should investigate and how to best integrate these individual differences into their study design.

Incorporating more explorations of individual differences into research with young bilinguals raises questions about what individual differences to incorporate. It is difficult and time-consuming to recruit large enough samples to look at individual differences and variation in bilingual language development. However, some studies have investigated individual variation in young bilinguals' language development. One commonly studied source of variation is the relative amount of exposure that a young bilingual hears in each language. For example, language exposure is strongly related to the size of their vocabulary in that language (Côté et al., 2022; Thordardottir, 2011). Another commonly studied source of variation is language similarity. For example, bilinguals learning similar languages (e.g., Dutch and Frisian) have been found to have larger vocabularies in Dutch than bilinguals learning dissimilar languages (e.g., Dutch and Polish; Blom et al., 2020). While relative language exposure and similarity are important components in bilinguals' language development (Byers-Heinlein, 2015), bilinguals' input and experiences can vary in many different ways that could impact their language development. Bilinguals' experiences likely vary in ways that have been shown to impact monolinguals' language development (e.g., total amount of input; Golinkoff et al., 2015; Shneidman & Goldin-Meadow, 2012), but they will also vary in ways unique to a bilingual experience (Hoff, 2020). For example, Chapter 4 showed that individual differences like the frequency children hear code-switching may impact their language processing and comprehension. To gain a more complete view of bilingualism, researchers will need to consider and identify which individual differences are the most relevant when investigating bilingual language development.

Beyond deciding which individual differences to incorporate, researchers must also determine how to best integrate these differences. An important consideration is that increasing variation on one variable can unintentionally increase variation on another. For instance, if researchers wanted to study the effect of age of exposure on language comprehension, the variation of age of exposure may also be related to variation in vocabulary size (e.g., Bylund et al., 2019), which in turn could affect language comprehension. Thus, it may be difficult to isolate the effect of age of exposure unless a large enough sample was collected to control for these other related variables. One way to address such complications would be for researchers to adopt a large-scale collaborative approach (see ManyBabies; Byers-Heinlein, Tsui, et al., 2021). This would allow for a larger and more diverse sample that could be used to assess the effects of individual differences by incorporating either the factor mixture model or grade-of-membership model, as proposed in Chapter 2. While investigating the potential effects of individual differences may complicate and lengthen the research process, it will ultimately provide a richer, more-detailed picture of bilingual language development.

5.2.3 Implications for Language Separation

The findings on code-switching in this dissertation also contribute to theories of bilingual language development. In particular, this work informs the question of language separation, or bilinguals' ability to treat the languages they hear as separate entities, which has been proposed to be a critical step in a bilingual child's language development (Byers-Heinlein, 2014). Code-switching offers a unique opportunity to explore language separation in young bilinguals, because it combines both languages in a short time period. This allows researchers to compare how bilinguals handle both of their languages simultaneously, as opposed to only one language at a time, which can provide information on how the languages are represented in the child's mind and if they are separated. For example, recognizing that a code-switch has occurred has been interpreted as evidence that bilinguals understand that separate languages were spoken (Byers-Heinlein et al., 2017). While code-switching offers a lens through which to investigate language separation in an experimental context, it is also relevant to ask how everyday exposure to code switching might impact a child's language development. Some scholars and parents have worried that code-switching may complicate language learning and separation, because it may reduce the contextual and linguistic cues that children can use to determine which language they are hearing (e.g., Baker, 2000; Barron-Hauwaert, 2004; Döpke, 1998; Taeschner, 1983). Below I explain what code-switching can reveal about language separation and argue that code-switching in the input may play a role in language separation by allowing young bilinguals to compare properties and features of both of their languages.

Language separation begins early in life but is a gradual process (Byers-Heinlein, 2014; Genesee, 1989), and code-switching may reveal details of this process. An early step in language separation is noticing that a switch between two languages has occurred. Chapter 3 found that for young French–English bilinguals in Montreal, the two languages appeared to be relatively separated in their language input. However, switches between the languages still occurred across larger time scales (e.g., across activities), and bilinguals would need to be able to detect when these switches happen. When switching happens on a larger time scale, bilinguals can detect such switches by using the rhythmic properties that distinguish the languages, an ability bilingual infants display within the first days and months of their life (Bosch & Sebastián-Gallés, 1997a, 1997b, 2001; Byers-Heinlein & Werker, 2009; Nazzi et al., 2000). However, when infants must detect switches between their languages on a smaller time scale due to code-switching, there may not be enough rhythmic information to accomplish this task. Thus, the ability to detect code-switching may be difficult particularly for young bilinguals. French–English bilinguals in Montreal have been shown to be unable to detect single-word switches between 8-12 months (Schott, Mastroberardino, et al., 2021), but they appear to develop this ability by 18-24 months (Byers-Heinlein et al., 2017). Chapter 4 further extends these findings and suggests that by 3 years old, French–English bilinguals in Montreal may have more refined abilities to handle their two language systems, as they were able to process code-switched sentences similar to single-language sentences. Combined, these findings suggest that the ability to detect code-switching emerges as young bilinguals gain more experience with their languages, supporting the idea of gradual language separation.

Code-switching not only reveals details of young bilinguals' language separation to researchers, but it may also reveal details of each language to the bilinguals themselves. Once young bilinguals are able to detect code-switching, code-switches may provide children with the opportunity to compare elements across their languages on a short time scale. For example, hearing the word "red" in English (i.e., /ɹɛd/) and French (i.e., "*rouge*," /ʁuʒ/) could allow a young bilingual to compare and learn the different way each language pronounces the r-sound. Being able to make these comparisons across languages has been proposed to be a key element in language separation (Curtin et al., 2011). It is through the comparison of linguistic features, such as phonotactics or lexical items, that children can learn which features belong to which language. When children can compare elements of their languages within the same sentence or between two adjacent sentences due to code-switching, it may highlight the differences across languages more clearly than when those same elements are compared across longer time scales.

While code-switching could help young bilinguals identify specific properties of each of their languages, it may be difficult to isolate this effect in the laboratory. Detecting code-switching and learning language-specific properties from code-switching could support each other. That is, detecting code-switching may help young bilinguals learn language-specific properties, and recognizing language-specific properties may help young bilinguals detect code-switching. Children's ability to detect code-switching appears to improve over time (Byers-Heinlein et al., 2017; Schott, Mastroberardino, et al., 2021), and the results from Chapter 3 suggest that parents increase the amount of code-switching they produce as their child's language abilities develop. If this is indeed the case, then it would not be surprising that older bilinguals, who would have more experience with language overall and code-switching specifically, would be able to detect code-switching, providing evidence for language separation, when younger bilinguals do not. One way to tease apart the potential contribution of code-switching to language separation would be to study children who hear different amounts of code-switching. If code-switching does in fact play a role in learning language-specific properties and language separation, children who hear code-switching frequently should be able to detect code-switching and language-specific properties more readily that children who hear code-switching infrequently. Understanding the impact of code-switching on language separation will shed light on a key step in bilingual language development.

5.3 Limitations & Future Directions

While this dissertation has addressed important issues surrounding code-switching in young bilinguals' language development and provided two new methodological contributions to the field of bilingualism, it also had several limitations, which raise additional questions that should be considered in future research.

One limitation was the use of a categorical definition of bilingualism used in Chapters 3 and 4. Based on the arguments presented in Chapter 2, it may seem counterintuitive that more individual differences were not considered in the subsequent chapters (although individual differences were incorporated to the extent possible, e.g., age, testing location, SES). This is largely because the designs of the studies presented Chapters 3 and 4, including the definition of bilingualism and sample sizes, were finalized before the ideas in Chapter 2 had crystalized. Additionally, research with young bilinguals is extremely time consuming, so the sample sizes tend to be small for studies in this field. This limits the incorporation of either the factor mixture model or grade-of-membership model proposed in Chapter 2. To illustrate, the factor mixture model or grade-of-membership model require a sample size of 150-200 participants to construct an accurate model (Holmes Finch, 2020; Lubke & Neale, 2006), but Chapter 3 had a final sample size of 21 families, and Chapter 4 had a final sample size of 30 children. In a world with unlimited resources and where the order of the work presented in this dissertation had been reversed, the ideas from Chapter 2 would have been better incorporated by increasing the sample size, expanding the definition of bilingualism, and identifying relevant individual differences to investigate in Chapters 3 and 4. This would have allowed for more subtle effects of bilingualism and code-switching to be explored.

Another limitation is the scope of Chapter 3. Chapter 3 focused on the code-switches that parents produced when speaking to their infant but was unable to look at how this code-switching may have directly impacted children's language development. This was because the corpus that was used to analyze parents' speech did not contain information about children's language abilities. Thus, future research could look at this link directly by investigating how the frequency, syntactic location, and/or apparent reason(s) behind parents' code-switching impacts their child's language development, including vocabulary size and comprehension skills. Additionally, the scope of Chapter 3 was limited to investigating the code-switching that infants heard from their parents, as the audio recordings were made at home. In a bilingual community like Montreal, it is very likely that infants were exposed to code-switching outside the home and/or from other speakers. While investigating the code-switching in a young bilingual's input more holistically could present some important privacy concerns, it would ultimately lead to a richer understanding of a bilingual's language environment. Lastly, the coding protocol used in Chapter 3 did not code for what part of speech was code-switched in intrasentential code-switches. As discussed in Chapter 4, gaining this more fine-grained information would allow researchers to examine how young bilinguals' familiarity with intrasentential code-switching at different parts of speech affects their language processing and comprehension.

The last limitations are related to the design and sample size in Chapter 4. Chapter 4 presented code-switching at a previously unstudied syntactic location – a prenominal determiner-adjective pair. Moreover, the adjectives were uninformative and did not differentiate between the two objects presented on the screen. This presented children with the opportunity to "listen through" the code-switch. Thus, it is unclear if children would be able to process all code-switching at an adjective as easily as they appeared to in this study. Thus, future research is needed to better understand bilinguals' comprehension of code-switching at a variety of syntactic locations. The other limitation present in Chapter 4 was the sample size. The sample afforded only a preliminary investigation of how children's individual differences may have affected their ability to process the code-switch. Moving forward, studies investigating bilinguals' comprehension of code-switching should plan on analyzing how individual differences affects bilinguals' performance and accommodate those analyses with an appropriate sample size.

5.4 Conclusion

Overall, this dissertation reveals two main patterns within the field of bilingualism. First, that there are more comprehensive ways to model bilingualism than are currently being used. If such models are used consistently in the field, they could support the comparison of results across studies and help address currently unanswered research questions, particularly around the effects of individual differences between bilinguals. Second, that there remains much to learn about the code-switching that young bilinguals encounter in their daily lives and how such code-switching may affect their language processing, language separation, and overall development. Results from this dissertation suggest that code-switching may support successful bilingual language development and that bilinguals' individual differences may moderate this effect.

References

Abutalebi, J., Brambati, S. M., Annoni, J.-M., Moro, A., Cappa, S. F., & Perani, D. (2007). The Neural Cost of the Auditory Perception of Language Switches: An Event-Related Functional Magnetic Resonance Imaging Study in Bilinguals. *Journal of Neuroscience*, *27*(50), 13762–13769. https://doi.org/10.1523/JNEUROSCI.3294-07.2007

Abutalebi, J., & Rietbergen, M. J. (2014). Neuroplasticity of the bilingual brain: Cognitive control and reserve. *Applied Psycholinguistics*, *35*(5), 895–899.

Albin, D. D., & Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behavior and Development*, *19*(4), 401–418. https://doi.org/10.1016/S0163-6383(96)90002-8

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Andreotti, A., Minicuci, N., Kowal, P., & Chatterji, S. (2009). Multidimensional Profiles of Health Status: An Application of the Grade of Membership Model to the World Health Survey. *PLOS ONE*, *4*(2), e4426. https://doi.org/10.1371/journal.pone.0004426

Antoniou, M. (2019). The Advantages of Bilingualism Debate. *Annual Review of Linguistics*, *5*(1), null. https://doi.org/10.1146/annurev-linguistics-011718-011820

Asparouhov, T., & Muthen, B. (2006). Multilevel modeling of complex survey data. *Proceedings of the Joint Statistical Meeting in Seattle*, 2718–2726.

Azuma, S. (1992). *Processing and intrasentential codeswitching* [PhD Thesis]. University of Texas, Austin.

Azuma, S. (2009). The frame-content hypothesis in speech production: Evidence from intrasentential code switching. *Linguistics*, *31*(6), 1071–1094. https://doi.org/10.1515/ling.1993.31.6.1071

Bail, A., Morini, G., & Newman, R. S. (2015). Look at the gato! Code-switching in speech to toddlers*. *Journal of Child Language*, *42*(5), 1073–1101. https://doi.org/10.1017/S0305000914000695

Bailey, B. (2000). Social/interactional functions of code switching among Dominican Americans. *Pragmatics*, *10*(2), 165–193. https://doi.org/10.1075/prag.10.2.01bai

Baker, C. (2014). *A Parents' and Teachers' Guide to Bilingualism*. Multilingual Matters.

Baker, W., & Trofimovich, P. (2005). Interaction of Native- and Second-Language Vowel System(s) in Early and Late Bilinguals. *Language and Speech*, *48*(1), 1–27. https://doi.org/10.1177/00238309050480010101

Barron-Hauwaert, S. (2004). *Language strategies for bilingual families*. Multilingual Matters.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics; New York*, *35*(5), 857–894. http://dx.doi.org.lib-ezproxy.concordia.ca/10.1017/S0142716414000174

Beatty-Martínez, A. L., Valdés Kroff, J. R., & Dussias, P. E. (2018). From the Field to the Lab: A Converging Methods Approach to the Study of Codeswitching. *Languages*, *3*(2), 19. https://doi.org/10.3390/languages3020019

Belazi, H. M., Rubin, E. J., & Toribio, A. J. (1994). Code Switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry*, *25*(2), 221–237. JSTOR.

Belsky, J. (1980). Mother-Infant Interaction at Home and in the Laboratory: A Comparative Study. *The Journal of Genetic Psychology*, *137*(1), 37–47. https://doi.org/10.1080/00221325.1980.10532800

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, *22*(1), e12715. https://doi.org/10.1111/desc.12715

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

Bialystok, E. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging*, *19*(2), 290–303. https://psycnet.apa.org/doi/10.1037/0882-7974.19.2.290

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological Validity in Bilingualism Research and the Bilingual Advantage. *Trends in Cognitive Sciences*, *22*(12), 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001

Blom, E., Boerma, T., Bosma, E., Cornips, L., van den Heuij, K., & Timmermeister, M. (2020). Cross-language distance influences receptive vocabulary outcomes of bilingual children. *First Language*, *40*(2), 151–171. https://doi.org/10.1177/0142723719892794

Blom, J.-P., & Gumperz, J. (1972). Social meaning in linguistic structures: Code-switching in Norway. In *Directions in Sociolinguistics* (pp. 407–434). Holt, Rinehart and Winston.

Bornstein, M. H., Gini, M., Putnick, D. L., Haynes, O. M., Painter, K. M., & Suwalsky, and J. T. D. (2006). Short-Term Reliability and Continuity of Emotional Availability in Mother–Child Dyads Across Contexts of Observation. *Infancy*, *10*(1), 1–16. https://doi.org/10.1207/s15327078in1001_1

Borsboom, D., Rhemtulla, M., Cramer, A. O. J., van der Maas, H. L. J., Scheffer, M., & Dolan, C. V. (2016). Kinds versus continua: A review of psychometric approaches to uncover the structure of psychiatric constructs. *Psychological Medicine*, *46*(8), 1567–1579. https://doi.org/10.1017/S0033291715001944

Bosch, L., & Sebastián-Gallés, N. (1997a). The role of prosody in infants' native-language discrimination abilities: The case of two phonologically close languages. *Fifth European Conference on Speech Communication and Technology*, 231–234. http://www.mirlab.org/conference_papers/International_Conference/Eurospeech%201997/pdf/mab/a0953.pdf

Bosch, L., & Sebastián-Gallés, N. (1997b). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69. https://doi.org/10.1016/S0010-0277(97)00040-1

Bosch, L., & Sebastián-Gallés, 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

Boudreault, M.-C., Cabirol, É.-A., Trudeau, N., Poulin-Dubois, D., & Sutton, A. (2007). Les inventaires MacArthur du développement de la communication: Validité et données normatives préliminaires [MacArthur Communicative Development Inventories: Validity and preliminary normative data]. *Canadian Journal of Speech-Language Pathology and Audiology*, *31*(1), 27–37.

Brown, D. J., Arnold, R., Standage, M., & Fletcher, D. (2017). Thriving on Pressure: A Factor Mixture Analysis of Sport Performers' Responses to Competitive Encounters. *Journal of Sport and Exercise Psychology*, *39*(6), 423–437. https://doi.org/10.1123/jsep.2016-0293

Bultena, S., Dijkstra, T., & Van Hell, J. G. (2015). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism*, *18*(3), 453. https://doi.org/10.1017/S1366728914000145

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), 455–474. https://doi.org/10.1017/S0142716407070257

Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism; Cambridge*, *16*(1), 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. (2015). Methods for studying infant bilingualism. In J. W. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (pp. 133–154). Cambridge University Press. https://doi.org/10.1017/CBO9781107447257.005

Byers-Heinlein, K., Bergmann, C., & Savalei, V. (2021). *Six solutions for more reliable infant research*. PsyArXiv. https://doi.org/10.31234/osf.io/u37fy

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%2F0956797609360758

Byers-Heinlein, K., Jardak, A., Fourakis, E., & Lew-Williams, C. (2021). Effects of language mixing on bilingual children's word learning. *Bilingualism: Language and Cognition*, 1–15. 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*, *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. (2020). MAPLE: A multilingual approach to parent language estimates. *Bilingualism: Language and Cognition*, *23*(5), 951–957. https://doi.org/10.1017/S1366728919000282

Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., Durrant, S., Fennell, C. T., Fiévet, A.-C., & Frank, M. C. (2021). A Multilab Study of Bilingual Infants: Exploring the Preference for Infant-Directed Speech. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920974622.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*(5), 815–823. https://doi.org/10.1111/j.1467-7687.2009.00902.x

Bylund, E., Abrahamsson, N., Hyltenstam, K., & Norrman, G. (2019). Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition*, *182*, 45–49. https://doi.org/10.1016/j.cognition.2018.08.020

Cantone, K. F., & MacSwan, J. (2009). The syntax of DP-internal codeswitching. *Multidisciplinary Approaches to Codeswitching*, 243–278.

Carbajal, M. J., & Peperkamp, S. (2020). Dual language input and the impact of language separation on early lexical development. *Infancy*, *25*(1), 22–45. https://doi.org/10.1111/infa.12315

Cheng, L.-R., & Butler, K. (1989). Code-switching: A natural phenomenon vs language 'deficiency.' *World Englishes*, *8*(3), 293–309. https://doi.org/10.1111/j.1467-971X.1989.tb00670.x

Clark, S. L., Muthén, B., Kaprio, J., D'Onofrio, B. M., Viken, R., & Rose, R. J. (2013). Models and Strategies for Factor Mixture Analysis: An Example Concerning the Structure Underlying Psychological Disorders. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 681–703. https://doi.org/10.1080/10705511.2013.824786

Cohen, J. (1983). The Cost of Dichotomization. *Applied Psychological Measurement*, *7*(3), 249–253. https://doi.org/10.1177/014662168300700301

Comeau, L., Genesee, F., & Lapaquette, L. (2003). The Modeling Hypothesis and child bilingual codemixing. *International Journal of Bilingualism*, *7*(2), 113–126. https://doi.org/10.1177/13670069030070020101

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Côté, S. L., Gonzalez-Barrero, A. M., & Byers-Heinlein, K. (2022). Multilingual toddlers' vocabulary development in two languages: Comparing bilinguals and trilinguals. *Journal of Child Language*, *49*(1), 114–130. https://doi.org/10.1017/S030500092000077X

Cummins, J. P. (1976). *The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. Working papers on bilingualism no. 9*. Bilingual Education Project, Ontario Institute for Studies in Education.

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

David, D. A., & Wei, L. (2008). Individual Differences in the Lexical Development of French–English Bilingual Children. *International Journal of Bilingual Education and Bilingualism*, *11*(5), 598–618. https://doi.org/10.1080/13670050802149200

de Bruin, A. (2019). Not All Bilinguals Are the Same: A Call for More Detailed Assessments and Descriptions of Bilingual Experiences. *Behavioral Sciences*, *9*(3), 33. https://doi.org/10.3390/bs9030033

De Cat, C., Gusnanto, A., & Serratrice, L. (2018). Identifying a threshold for the executive function advantage in bilingual children. *Studies in Second Language Acquisition*, *40*(1), 119–151. https://doi.org/10.1017/S0272263116000486

De Cat, C., Kašćelan, D., Prevost, P., Serratrice, L., Tuller, L., & Unsworth, S. (2021). *Delphi consensus survey on how to document bilingual experience*. PsyArXiv. https://doi.org/10.31219/osf.io/ebh3c

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, *2*(2011), 221–240. https://doi.org/10.1515/9783110239331.221

De Houwer, A., & Bornstein, M. H. (2016). Bilingual mothers' language choice in child-directed speech: Continuity and change. *Journal of Multilingual and Multicultural Development*, *37*(7), 680–693. https://doi.org/10.1080/01434632.2015.1127929

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 : JSLHR*, *59*(6), 1346–1356. https://doi.org/10.1044/2016_JSLHR-L-15-0234

Del Maschio, N., & Abutalebi, J. (2019). Language organization in the bilingual and multilingual brain. *The Handbook of the Neuroscience of Multilingualism*, 199–213. https://doi.org/10.1002/9781119387725.ch9

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

Dewaele, J.-M. (2010). *Emotions in Multiple Languages*. Springer.

Dewaele, J.-M., & Li, W. (2014). Intra- and inter-individual variation in self-reported code-switching patterns of adult multilinguals. *International Journal of Multilingualism*, *11*(2), 225–246. https://doi.org/10.1080/14790718.2013.878347

Dewaele, J.-M., & Zeckel, I. (2016). The psychological and linguistic profiles of self-reported code-switchers. *International Journal of Bilingualism*, *20*(5), 594–610. https://doi.org/10.1177/1367006915575411

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, *61*(12), 1203–1218. https://doi.org/10.1016/j.jbusres.2008.01.009

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Practical Assessment, Research, and Evaluation*, *14*(1). https://doi.org/10.7275/DA8T-4G52

Döpke, S. (1998). Can the principle of ‘one person – one language’ be disregarded as unrealisttcally elitist? *Australian Review of Applied Linguistics*, *21*(1), 41–56. https://doi.org/10.1075/aral.21.1.03dop

Dunn, A. L., & Fox Tree, J. E. (2009). A Quick, Gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition*, *12*(3), 273–289. https://doi.org/10.1017/S1366728909990113

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition*. https://doi.org/10.1037/t15144-000

Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). *Échelle de vocabulaire en images Peabody: EVIP*. Psycan.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *TVIP: Test de vocabulario en imágenes Peabody: adaptación Hispanoamericana*. American Guidance Service.

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: Parents simplify the structure of their speech in response to babbling. *Journal of Child Language*, *46*(5), 998–1011. https://doi.org/10.1017/S0305000919000291

Erosheva, E. A. (2005). Comparing Latent Structures of the Grade of Membership, Rasch, and Latent Class Models. *Psychometrika*, *70*(4), 619–628. https://doi.org/10.1007/s11336-001-0899-y

Erosheva, E. A., Fienberg, S. E., & Joutard, C. (2007). Describing disability through individual-level mixture models for multivariate binary data. *The Annals of Applied Statistics*, *1*(2), 346–384. https://doi.org/10.1214/07-aoas126

Esposito, A., Luk, G., Li, P., Tong, X., & Náñez, J. (2015). *Describing and quantifying "bilingualism": The need for consistency and accuracy for research and education*. Society for Research in Child Development, Philadelphia, PA.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories (CDIs)* (2nd ed.). Paul H. Brookes Publishing Company Baltimore, MD.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2014). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. PB Brookes.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 1497–1510. https://doi.org/10.2307/1130938

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Fernald, A., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in recognizing and understanding words in fluent speech. In *Approaches to Bootstrapping:*

Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Volume 1 (pp. 97–123). Benjamins.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*(4), 1003–1015. https://doi.org/10.1111/1467-8624.00331

Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive Psychology*, *60*(3), 190–217. https://doi.org/10.1016/j.cogpsych.2009.12.002

Fernandez, C. B., Litcofsky, K. A., & van Hell, J. G. (2019). Neural correlates of intra-sentential code-switching in the auditory modality. *Journal of Neurolinguistics*, *51*, 17–41. https://doi.org/10.1016/j.jneuroling.2018.10.004

Floor, P., & Akhtar, N. (2006). Can 18-Month-Old Infants Learn Words by Listening In on Conversations? *Infancy*, *9*(3), 327–339. https://doi.org/10.1207/s15327078in0903_4

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

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language*, *16*(01), 161–179. https://doi.org/10.1017/S0305000900013490

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*(3), 611–631. https://doi.org/10.1017/S0305000900009971

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, *4*(1), 1–6. https://doi.org/10.1038/ncomms2430

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to Me: The Social Context of Infant-Directed Speech and Its Effects on Early Language Acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Golinkoff, R. M., Hirsh-Pasek, K., & Schweisguth, M. (2001). A reappraisal of young children's knowledge of grammatical morphemes. *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, *1*, 167.

Gonzales, K., Byers-Heinlein, K., & Lotto, A. J. (2019). How bilinguals perceive speech depends on which language they think they're hearing. *Cognition*, *182*, 318–330. https://doi.org/10.1016/j.cognition.2018.08.021

Goodz, N. S. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal*, *10*(1), 25–44. https://doi.org/10.1002/1097-0355(198921)10:1<25::AID-IMHJ2280100104>3.0.CO;2-R

Gormley, I. C., & Murphy, T. B. (2009). A grade of membership model for rank data. *Bayesian Analysis*, *4*(2), 265–295. https://doi.org/10.1214/09-BA410

Gosselin, L., & Sabourin, L. (2021). Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia*, *159*, 107922. https://doi.org/10.1016/j.neuropsychologia.2021.107922

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(02), 67. https://doi.org/10.1017/S1366728998000133

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

Grosjean, F. (2001). The bilingual's language modes. In *One mind, two languages: Bilingual language processing* (pp. 1–22). Blackwell Pub.

Gullifer, J. W., & Titone, D. (2019). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 1–12. https://doi.org/10.1017/S1366728919000026

Gumperz, J. J. (1982). *Discourse Strategies* (Vol. 1). Cambridge University Press.

Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, *89*, 138–161. https://doi.org/10.1016/j.jml.2015.12.002

Hallquist, M. N., & Wright, A. G. C. (2014). Mixture Modeling Methods for the Assessment of Normal and Abnormal Personality, Part I: Cross-Sectional Models. *Journal of Personality Assessment*, *96*(3), 256–268. https://doi.org/10.1080/00223891.2013.845201

Heller, M. (2010). *Codeswitching: Anthropological and Sociolinguistic Perspectives* (Vol. 48). Walter de Gruyter.

Heredia, R. R., & Altarriba, J. (2001). Bilingual Language Mixing: Why Do Bilinguals Code-Switch? *Current Directions in Psychological Science*, *10*(5), 164–168. https://doi.org/10.1111/1467-8721.00140

Hills, T. (2013). The company that words keep: Comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language*, *40*(3), 586–604.

Hoff, E. (2020). Lessons from the study of input effects on bilingual development. *International Journal of Bilingualism*, *24*(1), 82–88. https://doi.org/10.1177/1367006918768370

Hoff, E., & Core, C. (2015). What Clinicians Need to Know about Bilingual Development. *Seminars in Speech and Language*, *36*(2), 89–99. https://doi.org/10.1055/s-0035-1549104

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development*. *Journal of Child Language*, *39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Holmes Finch, W. (2020). Performance of the Grade of Membership Model Under a Variety of Sample Sizes, Group Size Ratios, and Differential Group Response Probabilities for Dichotomous Indicators. *Educational and Psychological Measurement*, *81*(3), 523–548. https://doi.org/10.1177/0013164420957384

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective*. *Bilingualism: Language and Cognition*, *15*(2), 422–433. https://doi.org/10.1017/S1366728911000678

Ianco-Worrall, A. D. (1972). Bilingualism and Cognitive Development. *Child Development*, *43*(4), 1390–1400. https://doi.org/10.2307/1127524

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679. https://doi.org/10.1111/j.1467-7687.2008.00805.x

Jackson-Maldonado, D., Thal, D. J., & Fenson, L. (2003). *MacArthur Inventarios del Desarrollo de Habilidades Comunicativas: User's guide and technical manual*. Brookes Pub.

Kaushanskaya, M., & Crespo, K. (2019). Does Exposure to Code-Switching Influence Language Performance in Bilingual Children? *Child Development*, *90*(3), 708–718. https://doi.org/10.1111/cdev.13235

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710. https://doi.org/10.3758/PBR.16.4.705

Kaushanskaya, M., & Prior, A. (2015). Variability in the effects of bilingualism on cognition: It is not just about cognition, it is also about bilingualism. *Bilingualism: Language and Cognition*, *18*(1), 27–28. https://doi.org/10.1017/S1366728914000510

Kidd, E., & Garcia, R. (2021). *How diverse is child language acquisition?* PsyArXiv. https://doi.org/10.31234/osf.io/jpeyq

Kids Count Data Center. (2018). *The Number of Bilingual Kids in America Continues to Rise*. The Annie E. Casey Foundation: Kids Count Data Center. https://datacenter.kidscount.org/updates/show/184-the-number-of-bilingual-kids-in-america-continues-to-rise

Kircher, R. (2009). *Language attitudes in Quebec: A contemporary perspective* [PhD Thesis, Queen Mary University of London]. https://qmro.qmul.ac.uk/xmlui/handle/123456789/497

Kitamura, C., & Burnham, D. (2003). Pitch and Communicative Intent in Mother's Speech: Adjustments for Age and Sex in the First Year. *Infancy*, *4*(1), 85–110. https://doi.org/10.1207/S15327078IN0401_5

Kitamura, C., & Lam, C. (2009). Age-Specific Preferences for Infant-Directed Affective Intent. *Infancy*, *14*(1), 77–100. https://doi.org/10.1080/15250000802569777

Kosie, J. E. (2019). *Pupillometry as a Window on the Role of Motionese in Infants' Processing of Dynamic Activity* [PhD Thesis, University of Oregon]. www.proquest.com%2Fdissertations-theses%2Fpupillometry-as-window-on-role-motionese-infants%2Fdocview%2F2308259553%2Fse-2%3Faccountid%3D10246

Kroll, J. F., & Dussias, P. E. (2004). The comprehension of words and sentences in two languages. *The Handbook of Bilingualism*, 169–200.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In *Psychology of learning and motivation* (Vol. 56, pp. 229–262). Academic Press. https://doi.org/10.1016/B978-0-12-394393-4.00007-8

Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science*, *15*(2), 212–221. https://doi.org/10.1111/j.1467-7687.2011.01118.x

Lambert, W. E., Havelka, J., & Gardner, R. C. (1959). Linguistic Manifestations of Bilingualism. *The American Journal of Psychology*, *72*(1), 77–82. https://doi.org/10.2307/1420213

Law, K. S., Wong, C.-S., & Mobley, W. H. (1998). Toward a Taxonomy of Multidimensional Constructs. *The Academy of Management Review*, *23*(4), 741–755. https://doi.org/10.2307/259060

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394–425. https://doi.org/10.1037/bul0000142

Leyendecker, B., Lamb, M. E., & Schölmerich, A. (1997). Studying mother-infant interaction: The effects of context and length of observation in two subcultural groups. *Infant Behavior and Development*, *20*(3), 325–337. https://doi.org/10.1016/S0163-6383(97)90004-7

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods*, *38*(2), 202–210. https://doi.org/10.3758/BF03192770

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(03), 673–680. https://doi.org/10.1017/S1366728913000606

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2015). A Developmental Vocabulary Assessment for Parents (DVAP): Validating Parental Report of Vocabulary Size in 2- to 7-Year-Old Children. *Journal of Cognition and Development*, *16*(3), 442–454. https://doi.org/10.1080/15248372.2013.835312

Litcofsky, K. A., & Van Hell, J. G. (2017). Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia*, *97*, 112–139. https://doi.org/10.1016/j.neuropsychologia.2017.02.002

Lubke, G., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*(1), 21–39. https://doi.org/10.1037/1082-989X.10.1.21

Lubke, G., & Neale, M. C. (2006). Distinguishing Between Latent Classes and Continuous Factors: Resolution by Maximum Likelihood? *Multivariate Behavioral Research*, *41*(4), 499–532. https://doi.org/10.1207/s15327906mbr4104_4

Luk, G. (2015). Who are the bilinguals (and monolinguals)? *Bilingualism: Language and Cognition*, *18*(1), 35–36. https://doi.org/10.1017/S1366728914000625

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

Luk, G., Marian, V., Castro, D. C., & Byers-Heinlein, K. (2017). *Describing and quantifying "bilingualism" part 2; the need for consistency and accuracy for research and education*. Society for Research on Child Development, Austin, TX.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Macizo, P., Bajo, T., & Cruz Martín, M. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish–English interlexical homographs. *Journal of Memory and Language*, *63*(2), 232–244. https://doi.org/10.1016/j.jml.2010.04.002

Macnamara, J. (1967). The bilingual's linguistic performance—A psychological overview. *Journal of Social Issues*, *23*(2), 58–77.

MacSwan, J. (2012). Code-Switching and Grammatical Theory. In *The handbook of bilingualism and multilingualism* (p. 323).

MacWhinney, B. (2014). *The CHILDES project: Tools for analyzing talk, Volume II: The database*. Psychology Press.

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, *37*(4), 817–840. https://doi.org/10.1017/S0305000909990055

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2017). Caregiver talk to young Spanish-English bilinguals: Comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*, *20*(1), e12425. https://doi.org/10.1111/desc.12425

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and

Multilinguals. *Journal of Speech Language and Hearing Research*, *50*(4), 940. https://doi.org/10.1044/1092-4388(2007/067)

Marian, V., & Hayakawa, S. (2021). Measuring bilingualism: The quest for a "bilingualism quotient." *Applied Psycholinguistics*, *42*(2), 527–548. https://doi.org/10.1017/S0142716420000533

Marino, C., Bernard, C., & Gervain, J. (2020). Word Frequency Is a Cue to Lexical Category for 8-Month-Old Infants. *Current Biology*, *30*(8), 1380–1386. https://doi.org/10.1016/j.cub.2020.01.070

Masyn, K. E., Henderson, C. E., & Greenbaum, P. E. (2010). Exploring the Latent Structures of Psychological Constructs in Social Development Using the Dimensional–Categorical Spectrum. *Social Development*, *19*(3), 470–493. https://doi.org/10.1111/j.1467-9507.2009.00573.x

McLachlan, G., & Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

McNamara, T. F. (2000). *Language testing*. Oxford University Press.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, *50*(4), 266–275. https://doi.org/10.1037/0003-066X.50.4.266

Mirman, D. (2017). *Growth curve analysis and visualization using R*. Chapman and Hall/CRC.

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

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, *4*(1), 1–21. https://doi.org/10.1017/S1366728901000116

Muthén, L. K., & Muthén, B. (2016). Mplus. *The Comprehensive Modelling Program for Applied Researchers: User's Guide*, *5*.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.

Myers-Scotton, C. (2017). Code-switching. In *The handbook of sociolinguistics* (pp. 217–237).

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

Ng, S., Gonzalez, C., & Wicha, N. Y. Y. (2014). The fox and the cabra: An ERP analysis of reading code switched nouns and verbs in bilingual short stories. *Brain Research*, *1557*, 127–140. https://doi.org/10.1016/j.brainres.2014.02.009

Nilep, C. (2006). “Code Switching” in Sociocultural Linguistics. *Colorado Research in Linguistics*, *19*(1). https://doi.org/10.25810/hnq4-jv62

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

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*, *62*(7). 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

Orena, A. J., & Polka, L. (2019). Monolingual and bilingual infants’ word segmentation abilities in an inter-mixed dual-language task. *Infancy*, *24*(5), 718–737. https://doi.org/10.1111/infa.12296

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*, 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, *43*(6), 474–497. https://doi.org/10.1016/j.jcomdis.2010.01.002

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

Penfield, W., & Roberts, L. (1959). *Speech arid Brain-Mechanisms*. Princeton: Princeton U. Press.

Place, S., & Hoff, E. (2011). Properties of Dual Language Exposure That Influence 2-Year-Olds' Bilingual Proficiency. *Child Development*, *82*(6), 1834–1849. https://doi.org/10.1111/j.1467-8624.2011.01660.x

Place, S., & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds*. *Bilingualism: Language and Cognition*, *19*(5), 1023–1041. https://doi.org/10.1017/S1366728915000322

Pliatsikas, C., & Schweiter, J. W. (2019). Multilingualism and brain plasticity. In *The handbook of the neuroscience of multilingualism* (pp. 230–251). Wiley Online Library.

Polites, G. L., Roberts, N., & Thatcher, J. (2012). Conceptualizing models using multidimensional constructs: A review and guidelines for their use. *European Journal of Information Systems*, *21*(1), 22–48. https://doi.org/10.1057/ejis.2011.10

Poplack, S. (1978). *Syntactic structure and social function of code-switching* (Vol. 2). Centro de Estudios Puertorriqueños,[City University of New York].

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching1. *Linguistics*, *18*(7–8), 581–618. https://doi.org/10.1515/ling.1980.18.7-8.581

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. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*. 40th Annual Conference of the 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

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching*. *Bilingualism: Language and Cognition*, *13*(2), 253–262. https://doi.org/10.1017/S1366728909990526

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology*, *45*(2), 544. https://doi.org/10.1037/a0013917

Radford, A. (2006). *Minimalist Syntax Revisited*. http://courses.essex.ac.uk/lg/lg514

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, *17*(6), 880–891. https://doi.org/10.1111/desc.12172

Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, *53*(2), 818–835. https://doi.org/10.3758/s13428-020-01460-x

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, *55*(8), 1656–1665. https://doi.org/10.1037/dev0000754

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research*, *21*(4), 301–316. https://doi.org/10.1007/BF01067515

Ritchie, W. C., & Bhatia, T. K. (2012). Social and Psychological Factors in Language Mixing. *The Handbook of Bilingualism and Multilingualism*, 375.

Rocha-Hidalgo, J., & Barr, R. (2021). *Defining Bilingualism in Infancy and Toddlerhood: A Scoping Review*. PsyArXiv. https://doi.org/10.31234/osf.io/bmdhx

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated With Language-Related Brain Function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Sankoff, D., & Poplack, S. (1981). A formal grammar for code-switching. *Paper in Linguistics*, *14*(1), 3–45. https://doi.org/10.1080/08351818109370523

Schott, E., Kremin, L. V., & Byers-Heinlein, K. (2021). *The youngest bilingual Canadians: Insights from the 2016 Census about children aged 0-9*. PsyArXiv. 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

Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A Bilingual Advantage in Visual Language Discrimination in Infancy. *Psychological Science*, *23*(9), 994–999. https://doi.org/10.1177/0956797612436817

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659–673. https://doi.org/10.1111/j.1467-7687.2012.01168.x

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech*. *Bilingualism: Language and Cognition*, *16*(2), 304–324. https://doi.org/10.1017/S1366728912000466

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532. https://doi.org/10.1016/j.dr.2007.06.002

Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35*(4), 869–902. https://doi.org/10.1017/S0305000908008763

Soderstrom, M., Casillas, M., Bergelson, E., Rosemberg, C., Alam, F., Warlaumont, A. S., & Bunce, J. (2021). Developing a Cross-Cultural Annotation System and MetaCorpus for Studying Infants' Real World Language Experience. *Collabra: Psychology*, *7*(1), 23445. https://doi.org/10.1525/collabra.23445

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, *10*(1), 1–15. https://doi.org/10.1017/S0305000900005092

Stevenson, M. B., Leavitt, L. A., Roach, M. A., Chapman, R. S., & Miller, J. F. (1986). Mothers' speech to their 1-year-old infants in home and laboratory settings. *Journal of Psycholinguistic Research*, *15*(5), 451–461. https://doi.org/10.1007/BF01067725

Sulpizio, S., Del Maschio, N., Del Mauro, G., Fedeli, D., & Abutalebi, J. (2020). Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *NeuroImage*, *205*, 116306. https://doi.org/10.1016/j.neuroimage.2019.116306

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015*. *Bilingualism: Language and Cognition*, *22*(2), 401–415. https://doi.org/10.1017/S1366728917000682

Swingley, D. (2012). The Looking-While-Listening Procedure. In E. Hoff (Ed.), *Research Methods in Child Language* (pp. 29–42). Wiley-Blackwell. https://doi.org/10.1002/9781444344035.ch3

Taeschner, T. (1983). *The sun is feminine: A study on language acquisition in bilingual children*. Springer.

Takahesu Tabori, A. A., Mech, E. N., & Atagi, N. (2018). Exploiting Language Variation to Better Understand the Cognitive Consequences of Bilingualism. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01686

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6), e12456. https://doi.org/10.1111/desc.12456

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The Efficiency of Attentional Networks in Early and Late Bilinguals: The Role of Age of Acquisition. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00123

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*(4), 426–445. https://doi.org/10.1177/1367006911403202

Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition*, *100*(3), 389–433. https://doi.org/10.1016/j.cognition.2005.04.009

Tomić, A., & Kroff, J. R. V. (2021). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*, 1–12. https://doi.org/10.1017/S1366728921000237

Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: Evidence from eye-tracking. *Journal of Child Language;*

Cambridge, *43*(2), 425–441. http://dx.doi.org.lib-ezproxy.concordia.ca/10.1017/S0305000915000173

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender*. *Bilingualism: Language and Cognition*, *16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Dussias, P. E. (2018). Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches. *Linguistic Approaches to Bilingualism*, *8*(1), 98–133. https://doi.org/10.1075/lab.16011.val

van Hell, J. G., Fernandez, C. B., Kootstra, G. J., Litcofsky, K. A., & Ting, C. Y. (2018). Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism*, *8*(1), 134–161. https://doi.org/10.1075/lab.16010.van

Vaughan-Evans, A., Parafita Couto, M. C., Boutonnet, B., Hoshino, N., Webb-Davies, P., Deuchar, M., & Thierry, G. (2020). Switchmate! An Electrophysiological Attempt to Adjudicate Between Competing Accounts of Adjective-Noun Code-Switching. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.549762

Venker, C. E., Pomper, R., Mahr, T., Edwards, J., Saffran, J., & Weismer, S. E. (2020). Comparing Automatic Eye Tracking and Manual Gaze Coding Methods in Young Children with Autism Spectrum Disorder. *Autism Research*, *13*(2), 271–283. https://doi.org/10.1002/aur.2225

Viroli, C. (2012). *FactMixtAnalysis: Factor Mixture Analysis with covariates* (1.0) [Computer software]. https://CRAN.R-project.org/package=FactMixtAnalysis

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Sage Publications, Inc.

Wang, Y. S., & Erosheva, E. A. (2015). *mixedMem: Tools for Discrete Multivariate Mixed Membership Models* (1.1.0) [Computer software]. https://CRAN.R-project.org/package=mixedMem

Warlaumont, A. S., VanDam, M., Bergelson, E., & Cristia, A. (2017). HomeBank: A Repository for Long-Form Real-World Audio Recordings of Children. *INTERSPEECH*, 815–816.

Weinreich, U. (2010). *Languages in Contact: Findings and Problems*. Walter de Gruyter.

Werker, J. F., & Hensch, T. K. (2015). Critical Periods in Speech Perception: New Directions. *Annual Review of Psychology*, *66*(1), 173–196. https://doi.org/10.1146/annurev-psych-010814-015104

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

Woodbury, M. A., & Manton, K. G. (1989). Grade of Membership Analysis of Depression-Related Psychiatric Disorders. *Sociological Methods & Research*, *18*(1), 126–163. https://doi.org/10.1177/0049124189018001005

Woodcock, R. W., Muñoz-Sandoval, A. F., & Ruef, M. L. (2005). *Woodcock-Muñoz Language Survey—Revised*. 26.

Woolford, E. (1983). Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry*, *14*(3), 520–536. JSTOR.

Yacovone, A., Moya, E., & Snedeker, J. (2021). Unexpected words or unexpected languages? Two ERP effects of code-switching in naturalistic discourse. *Cognition*, *215*, 104814. https://doi.org/10.1016/j.cognition.2021.104814

Yip, V., Mai, Z., & Matthews, S. (2018). CHILDES for bilingualism. In *Bilingual Cognition and Language: The state of the science across its subfields* (Vol. 54, pp. 183–202). John Benjamins Publishing Company.

Yip, V., & Matthews, S. (2016). Code-mixing and mixed verbs in Cantonese-English bilingual children: Input and innovation. *Languages*, *1*(1), 4. https://doi.org/10.3390/languages1010004

Yow, W. Q., Tan, J. S. H., & Flynn, S. (2018). Code-switching as a marker of linguistic competence in bilingual children. *Bilingualism: Language and Cognition*, *21*(5), 1075–1090. https://doi.org/10.1017/S1366728917000335

Zeller, J. P. (2020). Code-Switching Does Not Equal Code-Switching. An Event-Related Potentials Study on Switching From L2 German to L1 Russian at Prepositions and Nouns. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01387

Zentella, A. C. (1981). *"Hablamos Los Dos. We Speak Both": Growing Up Bilingual in El Barrio.* [PhD Thesis, University of Pennsylvania]. https://repository.upenn.edu/dissertations/AAI8127100/

Zirnstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, *176*, 87–106. https://doi.org/10.1016/j.cognition.2018.03.001

Appendix A: Code-Switching Coding Instructions

Variable name	Class	Values	Description
r_row	Integer	Integers	Row number in R file
child_id	Integer	Integers	Unique child ID number
day	Integer	[1 2 3 4]	Day of recording
recording	Integer	1 – 1920	Number of recording in the day; unique 30 second chunk
speaker	Factor	Mom Dad Child Nanny	Who said the utterance
listener	Factor	Infant Mom Dad Child Nanny Other	Who the utterance was said to. We are only interested when the listener is "Infant"
language	Factor	English French Mixed	Which language utterance is spoken in. Mixed if more than one language spoken
transcription	String		Transcription of utterance "xxx" was unable to be transcribed "zzz" is redacted personal info, such as names
translation	String		Translation of entire utterance into English. For words that were originally in French, ARE IN ALL CAPS for easy identification.
multiple_switch	Factor	[0 1]	0 – if an utterance contains only one switch 1 – if an utterance contains more than one switch
switch_number	Integer	Integers	Any utterance that only has 1 switch will have a value of 1. For utterances with more than 1 switch, copy/paste the row for each switch. Each row will get its own number (2, 3, 4, etc.). *Bold and italicize* the switch in the translation column that is being coded in that row. Complete rest of row for switch of interest
direction	Factor	[1 2]	1 – if switch is from English to French 2 – if switch if from French to English
constituent	Factor	[1 2 3]	1 – if beginning of switch is not a constituent of pre-switch, and switch is between-sentences switches 2 – if beginning of switch is not a constituent, and switch is within-sentence 3 – if switch is a constituent[1] of pre-switch (only within-sentence switches possible)
attention	Logical	[0 1]	if switch is to attract infant's attention
emphasis	Logical	[0 1]	if switch is to add emphasis
discipline	Logical	[0 1]	if switch is to discipline the infant

[1] For description and tests of constituency: https://en.wikipedia.org/wiki/Constituent_(linguistics)

understanding	Logical	[0 1]	if switch is to ensure the infant's understanding
vocab	Logical	[0 1]	if switch is to teach vocabulary to the infant
translation_equivalent	Logical	[0 1]	if switch is translating the previous sentence resulting in an equivalent sentence
borrowing	Logical	[0 1]	if switch results from conventionalized, standard borrowing, such as a language specific saying, idiom, etc.
baby_word	Logical	[0 1]	If switch results from conventionalized, standard borrowing for baby vocabulary (e.g. dodo, toutou, bye-bye)
comments	String		If switch cannot be classified into a category above, add a comment if you can identify the reason. Other comments about the switch

Appendix B: Consent Forms

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 may take up to 60 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 data will be stored in a secure location at Concordia University, and participants will only be identified by code number to protect confidentiality.

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: _________________________

STANDARD ADULT CONSENT FORM WITH RECORDING
PRINCETON UNIVERSITY

Department of Psychology

TITLE OF RESEARCH: Language Learning: Sounds, Words, and Grammar

PRINCIPAL INVESTIGATOR: Casey Lew-Williams, Ph.D.

IRB PROTOCOL #: *7117*

IRB Approval Date: 8/12/19

IRB Expiration Date: 8/11/20

Your son/daughter is being invited to take part in a research study. Before you decide about your child's participation in this study, it is important that you understand why the research is being done and what it will involve. Please take the time to read the following information carefully. Please ask the researcher if there is anything that is not clear or if you need more information.

Purpose of Research: The purpose of this study is to better understand how children learn sounds, words, and grammar.

Procedures: This study will take 5-20 minutes, but appointments are scheduled for 45 minutes to allow time for full explanation of the purpose and procedure. You will be with your child the entire time in our research lab. Your child will listen to speech while viewing pictures or short animated movies. He/she may be tested to see what he/she learned. During the study, you will be asked to listen to music over headphones so that you do not learn what your child learns and accidentally influence his/her responses. We may also administer a standardized vocabulary test to your child, or other short questionnaires about child development. You may also be asked to complete a language development inventory at home before or after participating in this laboratory session. This inventory takes approximately 30 minutes to complete. You will receive monetary compensation ($10) for participating in this study, and your child will be rewarded with a small toy or book. You may also be compensated for travel expenses.

If you choose to allow your child to be in this study, your child has the right to be treated with respect, including respect for their decision of whether or not they with to continue or stop being in the study. Your child is free to stop being in the study at any time, even if you have already given permission for your child to be in the study. Choosing not to be in this study or to stop being in this study will not result in any penalty to your child or loss of benefits to which your child is otherwise entitled.

Video Recordings: Your child (and possibly you in the background) will be video-recorded during this procedure for research purposes, and laboratory staff will view the video at a later time. Video-recording is required for participation, and your child cannot participate if you do not agree to be recorded. Video recordings and photos of your child may be used for research purposes in the classroom, in publications, and at scientific meetings. Videos and any questionnaires will be kept in password-protected files that are only accessible to authorized researchers.

Revised: 7/20/18

Confidentiality: All aspects of your child's participation in this research will be confidential. We will not include any information that will make it possible to identify you or your child. Research records will be kept either in a locked cabinet or on a password-protected computer. Only authorized researchers will have access to your records. Any audio or video recordings will be destroyed after completion of the study. Results of this study may be used for teaching, research, publication, or presentations at scientific meetings. Your child's de-identified data file may also be shared with other research laboratories. If your child's individual results are discussed or shared with other research laboratories, his/her name and other identifying information will be protected using a code number rather than his/her name.

Risks or Discomforts/Benefits: There are no known physical or emotional risks involved in this study beyond those in everyday life.

Benefits: Your child is not likely to have any direct benefits from participating in this research study, but children usually enjoy the pictures/movies/sounds in the study. You will have the opportunity to learn more about how people learn language.

Who to contact with questions: If you have any questions about this study, please feel free to contact Professor Casey Lew-Williams. You can call him at telephone number (609) 258-7171, or send an email to caseylw@princeton.edu. If you have questions regarding your child's rights as a research subject, or if problems arise which you do not feel you can discuss with the Investigator, please contact the Institutional Review Board at: Office of Research Integrity and Assurance, Human Research Protection Program, (609) 258-3077, (609) 258-3105, irb@princeton.edu

---

I understand the following:

A. My child's participation in this study is voluntary, and I may withdraw my consent and discontinue participation at any time. My refusal to participate will not result in any penalty.

B. My child and I will be video- and audio-recorded during the study.

C. By signing this agreement, I do not waive any legal rights or release Princeton University, its agents, or you from liability for negligence.

I have read this form and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been told whom to contact. I agree to participate in the research study described above and will receive a copy of this consent form after I sign it. I certify that I am the legal parent (biological/ adoptive) and I hereby give my consent to be a subject in this research.

Parent's Name (printed)	Parent's Signature	Date

Child's Name

Revised: 7/20/18

Appendix C: Language Background Questionnaire

Baby ID: ____________ Exp. Name: ________________________

Study ID: ____________ Study Name: _______________________

Today's Date:___________________

Language Background Questionnaire For Children And Preschoolers					
Last Name		First Name		Today's Date	
Age		Date of Birth		Male ☐	Female ☐
(1) Please list all the languages your child knows in **order of dominance** (if languages are of the same dominance place them in the same box):					
1	2	3	4	5	
(2) Please list all the languages your child knows in **order of acquisition** (if your child acquired languages at the same time place them in the same box):					
1	2	3	4	5	
(3) Please list what percentage of the time your child is *on average* exposed to each language. (Percentages should add up to 100%):					
List language here:	A	B	C	D	E
List percentage here:					
(4) Please list what percentage of the time your child is *currently* exposed to each language. (Percentages should add up to 100%):					
List language here:	A	B	C	D	E
List percentage here:					
Language: English					
This is my child's	First ☐	Second ☐	Third ☐	Fourth ☐	Fifth ☐
All questions below refer to your child's knowledge of English					
(1) Age when your child...:					
first *heard* English:			was *regularly exposed* to English:		

(2) Please list the number of years and months your child spent in each language environment:					
			Years	Months	
A family where English is spoken					
A school and/or daycare where English is spoken					
(3) On a scale from zero to ten, please select your child's *level of proficiency* in speaking and understanding English:					
Speaking:			Understand spoken language:		
(4) On a scale from zero to ten, please rate to what extent your child is currently exposed to English in the following contexts:					
Interacting with family:		Watching TV:		Listening to the radio:	
Language: Spanish					
This is my child's	First ☐	Second ☐	Third ☐	Fourth ☐	Fifth ☐
All questions below refer to your child's knowledge of Spanish					
(1) Age when your child…:					
first *heard* Spanish:			was *regularly exposed* to Spanish:		
(2) Please list the number of years and months your child spent in each language environment:					
			Years	Months	
A family where Spanish is spoken					
A school and/or daycare where Spanish is spoken					
(3) On a scale from zero to ten, please select your child's *level of proficiency* in speaking and understanding Spanish:					
Speaking:			Understand spoken language:		
(4) On a scale from zero to ten, please rate to what extent your child is currently exposed to Spanish in the following contexts:					
Interacting with family:		Watching TV:		Listening to the radio:	
Language:					
This is my child's	First ☐	Second ☐	Third ☐	Fourth ☐	Fifth ☐

<table>
<tr><td colspan="6">All questions below refer to your child's knowledge of X</td></tr>
<tr><td colspan="6">(1) Age when your child…:</td></tr>
<tr><td colspan="3">first heard X:</td><td colspan="3">was regularly exposed to X:</td></tr>
<tr><td colspan="6">(2) Please list the number of years and months your child spent in each language environment:</td></tr>
<tr><td colspan="3"></td><td colspan="2">Years</td><td>Months</td></tr>
<tr><td colspan="3">A family where X is spoken</td><td colspan="2"></td><td></td></tr>
<tr><td colspan="3">A school and/or daycare where X is spoken</td><td colspan="2"></td><td></td></tr>
<tr><td colspan="6">(3) On a scale from zero to ten, please select your child's level of proficiency in speaking and understanding X:</td></tr>
<tr><td colspan="3">Speaking:</td><td colspan="3">Understand spoken language:</td></tr>
<tr><td colspan="6">(4) On a scale from zero to ten, please rate to what extent your child is currently exposed to X in the following contexts:</td></tr>
<tr><td colspan="2">Interacting with family:</td><td colspan="2">Watching TV:</td><td colspan="2">Listening to the radio:</td></tr>
<tr><td colspan="6">Language:</td></tr>
<tr><td>This is my child's</td><td>First ☐</td><td>Second ☐</td><td>Third ☐</td><td>Fourth ☐</td><td>Fifth ☐</td></tr>
<tr><td colspan="6">All questions below refer to your child's knowledge of X</td></tr>
<tr><td colspan="6">(1) Age when your child…:</td></tr>
<tr><td colspan="3">first heard X:</td><td colspan="3">was regularly exposed to X:</td></tr>
<tr><td colspan="6">(2) Please list the number of years and months your child spent in each language environment:</td></tr>
<tr><td colspan="3"></td><td colspan="2">Years</td><td>Months</td></tr>
<tr><td colspan="3">A family where X is spoken</td><td colspan="2"></td><td></td></tr>
<tr><td colspan="3">A school and/or daycare where X is spoken</td><td colspan="2"></td><td></td></tr>
<tr><td colspan="6">(3) On a scale from zero to ten, please select your child's level of proficiency in speaking and understanding X:</td></tr>
<tr><td colspan="3">Speaking:</td><td colspan="3">Understand spoken language:</td></tr>
<tr><td colspan="6">(4) On a scale from zero to ten, please rate to what extent your child is currently exposed to X in the following contexts:</td></tr>
<tr><td colspan="2">Interacting with family:</td><td colspan="2">Interacting with family:</td><td colspan="2">Interacting with family:</td></tr>
</table>

Appendix D: Developmental Vocabulary Assessment for Parents

For researcher use

Baby ID: ________________ Exp. name: __________________________

Study ID: ________________ Study name: _________________________

Today's date: _____________ Respondent's relationship to child: _________

Developmental Vocabulary Checklist for Parents

The aim of this questionnaire is to assess your child's current vocabulary. The list contains words that children learn between the ages of 2 and 18 years. We are interested in the words your child ***says***. Please mark each word you have heard your child say, including words that he/she pronounces differently (e.g. "doggie" instead of "dog") or that correspond to a different part of speech (e.g. "walked" instead of "walking"). If your child speaks more than one language, please only mark the words he/she says in English.

Boy	☐	Cookie	☐	Squirrel	☐	Buckle	☐
Chair	☐	Drum	☐	Throwing	☐	Sawing	☐
Puppy	☐	Turtle	☐	Farm	☐	Panda	☐
Bike	☐	Red	☐	Penguin	☐	Vest	☐
Laughing	☐	Jumping	☐	Gift	☐	Arrow	☐
Sleeping	☐	Carrot	☐	Feather	☐	Picking	☐
Hugging	☐	Reading	☐	Cobweb	☐	Target	☐
Walking	☐	Toe	☐	Elbow	☐	Dripping	☐
Ball	☐	Belt	☐	Juggling	☐	Knight	☐
Dog	☐	Fly	☐	Fountain	☐	Delivering	☐
Spoon	☐	Painting	☐	Net	☐	Cactus	☐
Foot	☐	Dancing	☐	Shoulder	☐	Dentist	☐
Duck	☐	Whistle	☐	Dressing	☐	Floating	☐
Banana	☐	Kicking	☐	Roof	☐	Claw	☐
Shoe	☐	Lamp	☐	Peeking	☐	Uniform	☐
Cup	☐	Square	☐	Ruler	☐	Gigantic	☐
Eating	☐	Fence	☐	Tunnel	☐	Furry	☐
Bus	☐	Empty	☐	Branch	☐	Violin	☐
Flower	☐	Happy	☐	Envelope	☐	Group	☐
Mouth	☐	Fire	☐	Diamond	☐	Globe	☐
Pencil	☐	Castle	☐	Calendar	☐	Vehicle	☐

Chef ☐
Squash ☐
Ax ☐
Flamingo ☐
Chimney ☐
Sorting ☐
Waist ☐
Vegetable ☐
Hyena ☐
Plumber ☐
River ☐
Timer ☐
Catching ☐
Trunk ☐
Vase ☐
Harp ☐
Bloom ☐
Horrified ☐
Swamp ☐
Heart ☐
Pigeon ☐
Ankle ☐
Flaming ☐
Wrench ☐
Aquarium ☐
Refueling ☐
Safe ☐
Boulder ☐
Reptile ☐
Canoe ☐
Athlete ☐
Towing ☐

Luggage ☐
Directing ☐
Vine ☐
Digital ☐
Dissecting ☐
Predatory ☐
Hydrant ☐
Surprised ☐
Palm ☐
Clarinet ☐
Valley ☐
Kiwi ☐
Interviewing ☐
Pastry ☐
Assisting ☐
Fragile ☐
Solo ☐
Snarling ☐
Puzzled ☐
Beverage ☐
Inflated ☐
Tusk ☐
Trumpet ☐
Rodent ☐
Inhaling ☐
Links ☐
Polluting ☐
Archaeologist ☐
Coast ☐
Injecting ☐
Fern ☐
Mammal ☐

Demolishing ☐
Isolation ☐
Clamp ☐
Dilapidated ☐
Pedestrian ☐
Interior ☐
Garment ☐
Departing ☐
Feline ☐
Hedge ☐
Citrus ☐
Florist ☐
Hovering ☐
Aquatic ☐
Reprimanding ☐
Carpenter ☐
Primate ☐
Glider ☐
Weary ☐
Hatchet ☐
Transparent ☐
Sedan ☐
Constrained ☐
Valve ☐
Parallelogram ☐
Pillar ☐
Consuming ☐
Currency ☐
Hazardous ☐
Pentagon ☐
Appliance ☐
Poultry ☐

Cornea ☐
Peninsula ☐
Porcelain ☐
Detonation ☐
Cerebral ☐
Perpendicular ☐
Submerging ☐
Syringe ☐
Lever ☐
Apparel ☐
Talon ☐
Cultivating ☐
Wedge ☐
Ascending ☐
Depleted ☐
Sternum ☐
Maritime ☐
Incarcerating ☐
Dejected ☐
Quintet ☐
Incandescent ☐
Confiding ☐
Mercantile ☐
Upholstery ☐
Filtration ☐
Replenishing ☐
Trajectory ☐
Perusing ☐
Barb ☐
Converging ☐
Honing ☐
Angler ☐

À l'usage du/ de la chercheur(se)	
Baby ID: ________________	Exp. name: ____________________________
Study ID: ________________	Study name: ___________________________
Today's date: _____________	Respondent's relationship to child: __________

Inventaire du développement du vocabulaire pour parents

L'objectif de ce questionnaire est de quantifier le vocabulaire de votre enfant. La liste est constituée de mots typiquement appris entre les âges de 2 et 18 ans. Nous sommes intéressés par les mots que votre enfant ***produit***. Veuillez s'il-vous-plaît cocher les mots que vous avez entendu votre enfant ***dire***. Ceci inclut les mots qu'il/elle prononce différemment (« wawal » pour « cheval ») ou qui occupe une différente fonction dans la phrase (« dormir » pour « dort »). Si votre enfant parle plus d'une langue, veuillez s'il-vous-plaît ne cocher que les mots qu'il/elle dit en français.

Poupée	☐	Bateau	☐	Coude	☐	Décoré	☐
Fourchette	☐	Autobus	☐	Bandage	☐	Tige	☐
Table	☐	Main	☐	Déchirer	☐	Tambourin	☐
Automobile	☐	Tracteur	☐	Forêt	☐	Repasseuse	☐
Homme	☐	Lit	☐	Mesurer	☐	Robinet	☐
Peigne	☐	Accident	☐	Enveloppe	☐	Voile	☐
Chaussette	☐	Tambour	☐	Hélicoptère	☐	Narine	☐
Bouche	☐	Vache	☐	Pneu	☐	Signal	☐
Se balancer	☐	Serpent	☐	Vide	☐	Surpris	☐
Boire	☐	Lampe	☐	Nid	☐	Groupe	☐
Marcher	☐	Genou	☐	Cage	☐	Remplir	☐
Grimper	☐	Plume	☐	Griffe	☐	Peler	☐
Roue	☐	Pingouin	☐	S'étirer	☐	Dispute	☐
Fermeture éclair	☐	Clôture	☐	Attacher	☐	Plonger	☐
Câble	☐	Parachute	☐	Flatter	☐	Livrer	☐
Râteau	☐	Flèche	☐	Coller	☐	Démolir	☐
Géant	☐	Carré	☐	Coudre	☐	Pot	☐
Mariée	☐	Filet	☐	Gonflé	☐	Écorce	☐
Sorcière	☐	Outil	☐	Épaule	☐	Dégoutter	☐
Royal	☐	Légume	☐	Cadre	☐	Balcon	☐

Hameçon	☐	Classer	☐	Belette	☐	Brindille	☐
Récompenser	☐	Véhicule	☐	Incertitude	☐	Inclément	☐
Fatigué	☐	Pyramide	☐	Serres	☐	Calice	☐
Cérémonie	☐	Isolement	☐	Ascension	☐	Émacié	☐
Mécanicien	☐	Délabré	☐	Charogne	☐	Spectre	☐
Fragile	☐	Médaillon	☐	Boulon	☐	Cornée	☐
Tronc	☐	Sommeiller	☐	Extenuée	☐	Entravé	☐
Anneau	☐	Ajustable	☐	Félin	☐	Enjoliver	☐
Vase	☐	Dromadaire	☐	Confidence	☐	Jubilante	☐
Tir à l'arc	☐	Extérieur	☐	Losange	☐	Mercantile	☐
Ustensile	☐	Reptile	☐	Arche	☐	Incandescent	☐
Casserole	☐	Trajectoire	☐	Constellation	☐	Obélisque	☐
Pédale	☐	Crâne	☐	Seringue	☐	Palan	☐
Colère	☐	Vigne	☐	Indigent	☐	Agrume	☐
Tranquillité	☐	Coopération	☐	Perpendiculaire	☐	Restreindre	☐
Cylindrique	☐	Penderie	☐	Assaillir	☐	Divergence	☐
Infirme	☐	Charpentier	☐	Arrogant	☐	Convexe	☐
Globe	☐	Nautique	☐	Péninsule	☐	Déambulation	☐
Expliquer	☐	Déception	☐	Spatule	☐	Larcin	☐
Disséquer	☐	Cascade	☐	Filtration	☐	Émission	☐
Humain	☐	Pelucheux	☐	Consommer	☐	Tangente	☐
Île	☐	Quatuor	☐	Aride	☐	Entomologiste	☐
Moulinet	☐	Vitrifié	☐	Défense	☐	Homoncule	☐
Transparent	☐	Avachi	☐	Côte	☐	Dénuement	☐
Communication	☐	Parallélogramme	☐	Abrasif	☐	Repoussé	☐
Piéton	☐	Cachet	☐	Urne	☐	Anthropoïde	☐
Enflammé	☐	Sphérique	☐	Solennel	☐		
Crampon	☐	Rembourrage	☐	Contempler	☐		

Subject ID: ________________________________

Today's Date:

Relationship to child: ____________________

Lista de Vocabulario del Desarrollo para Padres

El objetivo de este cuestionario es evaluar el vocabulario actual de su hijo/a. La lista incluye palabras que niños aprenden entre los edades de 2 a 18 años. Nos interesa las palabras que su hijo/a ***dice.*** Por favor, marque cada palabra que ha escuchado a su hijo/a decir, incluyendo palabras que pronuncia de manera diferente (por ejemplo, "pato" en vez de "zapato") o que corresponden a una categoría gramatical diferente (por ejemplo, "abre" en vez de "abriendo"). Si su hijo/a habla más que un idioma, por favor sólo marque las palabras que dice en español.

Muñeca ☐
Hombre ☐
Columpiar ☐
Tenedor ☐
Peine ☐
Beber ☐
Mesa ☐
Media ☐
Andar ☐
Perro ☐
Boca ☐
Subir ☐
Rueda ☐
Mapear ☐
Trapear ☐
Cierre ☐
Podar ☐
Soga ☐
Aserrar ☐
Rastrillo ☐
Pasear ☐
Barco ☐
Lámpara ☐
Vaca ☐
Vela ☐
Trompeta ☐
Rodilla ☐
Jaula ☐

Ambulancia ☐
Leer ☐
Flecha ☐
Cuello ☐
Mueble ☐
Abeja ☐
Hora ☐
Medir ☐
Ballena ☐
Roto ☐
Acariciar ☐
Accidente ☐
Canguro ☐
Codo ☐
Río ☐
Àguila ☐
Romper ☐
Pintor ☐
Vacío ☐
Pelar ☐
Uniforme ☐
Tronco ☐
Líquido ☐
Grupo ☐
Músico ☐
Ceremonia ☐
Culebra ☐
Bebida ☐

Médico ☐
Aislamiento ☐
Mecánico ☐
Premiar ☐
Dentista ☐
Hombro ☐
Sobre ☐
Joyas ☐
Humano ☐
Artista ☐
Recoger ☐
Construcción ☐
Dirigir ☐
Arbusto ☐
Bosque ☐
Agricultura ☐
Raíz ☐
Nutritivo ☐
Par ☐
Secretaria ☐
Iluminación ☐
Carrete ☐
Transparente ☐
Cosechar ☐
Discusión ☐
Cooperación ☐
Barandal ☐
Sorprendido ☐

Gotear	☐	Carroña	☐	Aparición	☐
Embudo	☐	Brújula	☐	Prodigio	☐
Tallo	☐	Esférico	☐	Ascender	☐
Isla	☐	Felino	☐	Fragmento	☐
Àngulo	☐	Paralelo	☐	Perpendicular	☐
Desilusión	☐	Sumergir	☐	Atuendo	☐
Carpintero	☐	Àrido	☐	Córnea	☐
Archivar	☐	Frágil	☐	Paralelogramo	☐
Mercantil	☐	Instruir	☐	Copioso	☐
Cuarteto	☐	Arqueólogo	☐	Inducir	☐
Marco	☐	Consumir	☐	Atónito	☐
Binocular	☐	Incandescente	☐	Transeúnte	☐
Judicial	☐	Arrogante	☐	Emisión	☐
Roer	☐	Utensilio	☐	Obelisco	☐
Morsa	☐	Ira	☐	Ciénaga	☐
Confiar	☐	Cítrico	☐	Ambulante	☐
Terno	☐	Lubricar	☐	Cóncavo	☐
Contemplar	☐	Eslabón	☐	Incisivo	☐
Ave	☐	Morada	☐	Elipse	☐
Portátil	☐	Anfibio	☐	Deciduo	☐
Clasificar	☐	Jubilosa	☐		

Appendix E: 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 <u>English</u> 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 <u>French</u> 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 <u>English</u> and then switch to speaking <u>French.</u>

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

e) I often start a sentence in <u>French</u> and then switch to speaking <u>English.</u>

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

f) I often borrow a <u>French</u> word when speaking <u>English.</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 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 F: Demographic Questionnaire

Today's Date: _____________ Baby ID: _______ Exp. Name: _________________

Study ID: ________ Study Name:___________________________

Concordia Infant Research Laboratory Participant Information

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

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

If yes, which university? __________________________________

Languages Spoken in the Home and at Childcare

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?) _________________________________

Health History

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 or vision problems**?

If yes please detail____

Does your child **currently** have an ear infection? ☐**Yes** ☐ **No**

Has your child had any ear infections **in the past**? ☐**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)?

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:

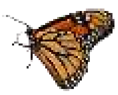
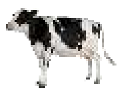
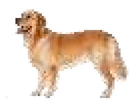
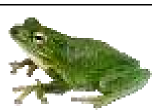
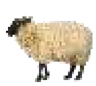
- ☐ Aboriginal
- ☐ African
- ☐ Arab
- ☐ West Asian
- ☐ South Asian
- ☐ East and Southeast Asian
- ☐ Caribbean
- ☐ European
- ☐ Latin/Central/South American
- ☐ Pacific Islands
- ☐ Not Applicable/Unknown
- ☐ Other (please specify): ____

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 G: Visual Stimuli for Chapter 4

Table G.1: Target images and their adjective-noun phrases in English, French, and Spanish.

Image	English	French	Spanish
	the good bear	-	*el buen oso*
	the big bunny the nice bunny	- *le gentil lapin*	*el gran conejo* -
	the little butterfly	-	*la pequeña mariposa*
	the pretty cow	*la jolie vache*	*la hermosa vaca*
	the big dog the nice dog	- *le gentil chien*	*el gran perro* -
	the good duck	*le bon canard*	*el buen pato*
	the good fish	*le bon poisson*	-
	the pretty froggy	*la jolie grenouille*	*la hermosa rana*
	the little monkey	*le petit singe*	-
	the little sheep	*le petit mouton*	*la pequeña oveja*

Note. Image pairs for each trial are described in Table 4.2.

Appendix H: Supplemental Analyses for Chapter 4

Growth Curve Analysis

As described in the main text, we conducted a growth curve analysis using the time window of 400 – 2000ms after the noun onset, with looking time data binned in 100ms blocks. The final model was built and selected through an iterative process (see Table H.1 for terms in each model; Mirman, 2017). First, we constructed a baseline model which contained only linear and quadratic time terms as fixed effects and random slopes on the random effect of participant (see Table H.2). For Model 1, we added the fixed effect of trial type, which was coded using a simple contrast coding scheme (see Table H.3). Compared to baseline, the addition of trial type significantly improved the model (see Table H.4), thus this variable was retained for subsequent models. For Model 2, we next added the interaction between trial type and the linear time term (see Table H.5). Compared to Model 1, the addition of this interaction did not statistically significantly improve the model (see Table H.6) and was removed for subsequent models. Model 1 served as the final model for our analysis on the effect of trial type and the baseline model for our analyses on the effects of language dominance, testing location, socioeconomic status, and vocabulary.

For each subsequent model, the variable of interest was added as a main effect and in an interaction with trial type. In Model 3, we investigated the effect of language dominance (see Table H.7). Compared to Model 1, the addition of language dominance did not significantly improve the model (see Table H.8). In Model 4, we investigated the effect of testing location (see Table H.9). Compared to Model 1, the addition of testing location statistically significantly improved the model (see Table H.10). In Model 5, we investigated the effect of socioeconomic status (see Table H.11). Compared to Model 1, the addition of socioeconomic status statistically significantly improved the model (see Table H.12). In Model 6, we investigated the effect of vocabulary (see Table H.13). Compared to Model 1, the addition of vocabulary statistically significantly improved the model (see Table H.14).

Pupillometry

For the sample collected in Montreal, we were able to examine whether hearing a code-switched sentence elicits a processing cost via pupillometry, as the Tobii T60-XL eye-tracker automatically records pupil size. Following the guidelines from Jackson and Sirois (2009), the data were pre-processed using the PupillometryR package (Forbes, 2020) in order to facilitate the analysis. The pre-processing began with regressing one pupil against the other. Because pupils often change size at a similar rate (Jackson & Sirois, 2009), this step allows data from one pupil to approximate the other when there is missing data. Then, the mean pupil size across both eyes was calculated in order to have a single pupil measure for each time sample. Data were then filtered with a moving hanning filter. The hanning filter calculates the pupil size through a weighted moving average to remove extreme values but keep the relevant effects (Kosie, 2019). Next, the data were baseline corrected by subtracting the average size of the pupil during the last 200ms of the carrier phrase before the first code-switch (e.g., 200ms before "*le bon*" in the sentence "Can you find *le bon* [fr. the good] duck?"). Baselining allows the change in pupil dilation within a window to be analyzed, as opposed to analyzing the raw pupil size which can drift over time. Trials were excluded from the pupillometry analyses if they were not analyzed in the looking time analysis (i.e., the child looked less than 750ms of the analysis window) or if they had no data during the baselining period or analysis windows. A total of 137 single-language trials and 120 code-switched trials were included in the following analyses (97% and 94% of single-language and code-switched trials included in the looking time analysis, respectively).

In order to isolate the effect of the code-switch from the carrier phrase to the article (e.g., Can you find *le bon* [fr. the good] …) and from the adjective to the noun (e.g., … *le bon* [fr. the good] duck?), we conducted separate analyses time locking the data at the location of each code-switch. First, we examined pupil dilation for single-language and code-switched trials for 2000ms from the onset of the article. Across this time window, pupil dilation was similar between the two trial types, $t(18) = 1.34$, $p = .198$, $M_d = 0.01$, 95% CI [-0.01,0.03], indicating no differences in processing effort for the code-switched trials immediately after the first switch (see Figure H.1a). Next, we examined pupil dilation for single-language and code-switched trials for 2000ms after the onset of the noun in the stimulus. Across this time window, pupil dilation was similar between the two trial types, $t(18) = 0.39$, $p = .700$, $M_d = 0.01$, 95% CI [-0.04,0.05], suggesting no difference in processing effort for the code-switched trials when the language switched again at the noun (see Figure H.1b).

Because articles began at different times across trials relative to the onset of the noun due to natural variation in speaking rate and length of the adjective (e.g., beautiful vs. old; M = 507ms, *Range*: 311 – 705ms), we extended the time window visualized in Figure H.1a to match the analysis window of the switch to the noun, which lasted until approximately 500ms after the analysis window for the first switch. Similarly, in Figure H.1b, we extended the window visualized to 500ms before the noun onset to approximately match the analysis window of the switch to the article.

Table H.1: Fixed effects in each model in the iterative process.

Model	Fixed effects
Base model	Linear time + quadratic time
Model 1	Linear time + quadratic time + trial type
Model 2	Linear time + quadratic time + trial type + trial type x linear time
Model 3	Linear time + quadratic time + trial type + language dominance + trial type x language dominance
Model 4	Linear time + quadratic time + trial type + testing location + trial type x testing location
Model 5	Linear time + quadratic time + trial type + socioeconomic status + trial type x socioeconomic status
Model 6	Linear time + quadratic time + trial type + vocabulary + trial type x vocabulary

Note. Each model had the same random effect structure including linear and quadratic time as random slope for participants.

Table H.2: Baseline Model.

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.76	[0.72, 0.79]	43.06	29.43	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.86	29.45	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.36	29.12	< .001
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.155			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + ([time (linear) + time (quadratic)] || participant)*

Table H.3: Model 1, addition of trial type as fixed factor.

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.76	[0.72, 0.79]	43.05	29.42	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.86	29.46	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.36	29.09	< .001
Trial type	-0.03	[-0.05, -0.01]	-3.43	6,100.82	.001
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.154			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + ([time (linear) + time (quadratic)] || participant)*

Table H.4: ANOVA comparison of Model 1 to the baseline model.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	X^2	*df*	*p*
Base model	7.00	6,043.90	6,090.99	-3,014.95	6,029.90			
Model 1	8.00	6,034.12	6,087.95	-3,009.06	6,018.12	11.78	1.00	0.00

Table H.5: Model 2, addition of interaction between trial type and time (linear).

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.76	[0.72, 0.79]	43.09	29.42	< .001
Time (Linear)	0.29	[0.14, 0.44]	3.85	29.44	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.37	29.07	< .001
Trial type	-0.03	[-0.05, -0.02]	-3.49	6,100.99	< .001
Trial type x Time (Linear)	-0.06	[-0.14, 0.02]	-1.46	6,103.57	.144
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.156			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + trial type:time (linear) + ([time (linear) + time (quadratic)] || participant)*

Table H.6: ANOVA comparison of Model 2 to Model 1.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	X^2	*df*	*p*
Model 1	8.00	6,034.12	6,087.95	-3,009.06	6,018.12			
Model 2	9.00	6,033.98	6,094.54	-3,007.99	6,015.98	2.14	1.00	0.14

Table H.7: Model 3, addition of language dominance.

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.76	[0.72, 0.79]	44.35	29.44	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.86	29.46	.001
Time (Quadratic)	-0.27	[-0.32, -0.23]	-12.37	29.08	< .001
Trial type	-0.03	[-0.05, -0.01]	-3.39	6,101.58	.001
Language dominance	-0.05	[-0.11, 0.02]	-1.36	29.44	.183
Trial type x Language dominance	0.01	[-0.03, 0.05]	0.35	6,101.58	.727
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.154			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + language dominance + trial type x language dominance ([time (linear) + time (quadratic)] || participant)*

Table H.8: ANOVA comparison of Model 3 to Model 1.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	X^2	*df*	*p*
Model 1	8.00	6,034.12	6,087.95	-3,009.06	6,018.12			
Model 3	10.00	6,036.19	6,103.48	-3,008.10	6,016.19	1.93	2.00	0.38

Table H.9: Model 4, addition of testing location.

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.75	[0.71, 0.78]	43.20	29.72	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.87	29.47	.001
Time (Quadratic)	-0.27	[-0.31, -0.23]	-12.44	29.05	< .001
Trial type	-0.05	[-0.07, -0.03]	-4.67	6,103.15	< .001
Testing location	-0.07	[-0.14, 0.00]	-1.98	29.71	.057
Testing location x Trial type	-0.09	[-0.13, -0.05]	-4.16	6,103.14	< .001
Random effects		**Variance**			
Participant	Intercept	0.007			
	Time (Linear)	0.154			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + testing location + testing location x trial type + ([time (linear) + time (quadratic)] || participant)*

Table H.10: ANOVA comparison of Model 4 to Model 1.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	X^2	*df*	*p*
Model 1	8.00	6,034.12	6,087.95	-3,009.06	6,018.12			
Model 4	10.00	6,017.25	6,084.53	-2,998.63	5,997.25	20.87	2.00	< .001

Table H.11: Model 5, addition of parental education (SES).

	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.69	[0.55, 0.83]	9.73	29.56	< .001
Time (Linear)	0.29	[0.14, 0.43]	3.85	29.47	.001
Time (Quadratic)	-0.27	[-0.31, -0.23]	-12.41	29.06	< .001
Trial type	-0.20	[-0.28, -0.12]	-4.75	6,106.01	< .001
Parental education	0.00	[0.00, 0.01]	0.94	29.48	.355
Testing location x Parental education	0.01	[0.01, 0.02]	4.04	6,103.72	< .001
Random effects		**Variance**			
Participant	Intercept	0.008			
	Time (Linear)	0.155			
	Time (Quadratic)	0.002			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + parental education + testing location x parental education + ([time (linear) + time (quadratic)] || participant)*

Table H.12: ANOVA comparison of Model 5 to Model 1.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	X^2	*df*	*p*
Model 1	8.00	6,034.12	6,087.95	-3,009.06	6,018.12			
Model 5	10.00	6,021.09	6,088.37	-3,000.54	6,001.09	17.03	2.00	< .001

Table H.13: Model 6, addition of vocabulary.

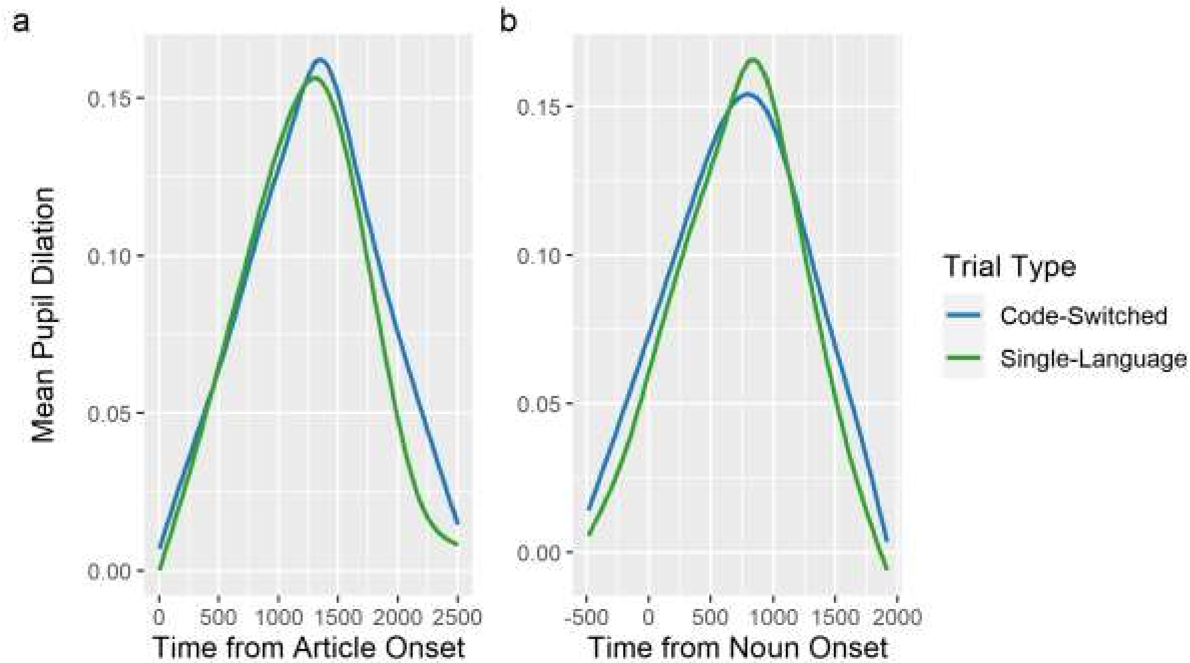
	Estimate	95% CI	t	df	p
Fixed effects					
Intercept	0.676	[0.6023,0.7497]	17.98	28.68	< .001
Time (Linear)	0.2619	[0.1208,0.4031]	3.64	28.35	.001
Time (Quadratic)	-0.2636	[-0.3050,-0.2222]	-12.48	28.16	< .001
Trial type	-0.0497	[-0.0960,-0.0034]	-2.1	5899.58	.035
Vocabulary	0.0007	[0.0001,0.0013]	2.42	28.38	.022
Trial type x Vocabulary	0.0002	[-0.0002,0.0005]	0.85	5896.3	.396
Random effects		**Variance**			
Participant	Intercept	0.007			
	Time (Linear)	0.138			
	Time (Quadratic)	0.001			

Note. Equation = *proportion looking time ~ [time (linear) + time (quadratic)] + trial type + vocabulary + testing location x vocabulary + ([time (linear) + time (quadratic)] || participant)*

Table H.14: ANOVA comparison of Model 6 to Model 1.

Model	Number of parameters	AIC	BIC	Log-likelihood	Deviance	*X*2	*df*	*p*
Model 1	8.00	5,863.09	5,916.64	-2,923.55	5,847.09			
Model 6	10.00	5,860.94	5,927.88	-2,920.47	5,840.94	6.15	2.00	0.05

Figure H.1: Mean change in pupil dilation by trial type from (a) article onset and (b) noun onset.



Note. We encourage the reader to interpret this figure with caution as the onsets of the nouns (in a) and onsets of articles (in b) do not all occur at the same time on this visualization.