**Growing up bilingual: examining the language input and
word segmentation abilities of bilingual infants**

Adriel John Orena

School of Communication Sciences and Disorders
McGill University, Montréal, QC
November 2018

A thesis submitted to McGill University in partial fulfillment of the requirements
of the degree of Doctor of Philosophy in Communication Sciences and Disorders.

Table of Contents

ABSTRACT

Infants' early language experiences play a critical role on their language development. In this dissertation, I explored the nature of this relationship in a bilingual context. Specifically, I investigated how bilingual caregivers are providing language input to their infants, and how global measures of this bilingual experience affect early word segmentation (i.e., the ability to recognize words in a sentential context). This work is important for understanding the factors that contribute to the developmental trajectory and processing capacities of bilingual infants.

In the first part of this dissertation, I assessed research methods for examining the language input to bilingual infants. To do so, I recruited twenty-one French-English bilingual families with a 10-month-old infant from Montréal, Canada. These families completed language interviews and contributed three full-day recordings at home using the LENA (**L**anguage **En**vironment **A**nalysis) recording system. *Chapter 2* provides support for using the LENA recording system for investigating the language input in bilingual infants, and *Chapter 3* shows that caregivers are reliable at describing their infants' language experience at home.

Next, I described the variability in language experiences within bilingual infants, and how these language experiences might affect word segmentation. In *Chapter 4,* I recruited 8- and 10-month-old infants from monolingual and bilingual homes. Our findings confirm that monolingual infants can segment bisyllabic words in their native language, but not a non-native language. Critically, our findings reveal that some bilingual infants are able to segment bisyllabic words in both of their native languages by 8-months of age.

Interestingly, exploratory analyses suggest that infants' word segmentation skills in our dual-language task are bolstered if they hear more language mixing from their caregivers.

In sum, this dissertation contributes to the growing literature that highlights the wide variability in bilingual language experiences, and their effects on early speech processing skills. Indeed, examining the language experiences and skills of bilingual-learning infants provides us with a unique lens for investigating language acquisition and development.

RÉSUMÉ

Les expériences linguistiques précoces des bébés jouent un rôle important dans leur développement langagier. Dans cette dissertation, j'ai exploré la nature de cette relation dans un contexte bilingue. Plus précisément, j'ai étudié comment les parents bilingues fournissent l'apport langagier à leurs bébés, et comment des mesures globales de cette exposition bilingue influencent la segmentation précoce des mots (c.-à-d. la capacité de reconnaitre des formes de mots dans le contexte de phrases). Les présents travaux sont importants afin de mieux comprendre les facteurs qui contribuent à la trajectoire développementale et aux capacités de traitement des bébés bilingues.

Dans la première partie de cette dissertation, j'ai comparé des méthodes de recherche servant à l'observation de l'apport langagier chez les bébés bilingues. Pour ce faire, j'ai recruté vingt-et-une familles bilingues en français et en anglais de Montréal, Canada, ayant un bébé de 10 mois. Ces familles ont pris part à des entrevues sur le langage et ont contribué trois journées entières d'enregistrements à la maison en utilisant le système d'enregistrement LENA ("**L**anguage **En**vironment **A**nalysis", ou analyse de l'environnement langagier). Le *Chapitre 2* soutient l'utilisation du système d'enregistrement LENA pour l'étude de l'apport langagier chez les bébés bilingues, et le *Chapitre 3* démontre que les parents décrivent l'exposition langagière de leurs enfants à la maison de façon juste.

Ensuite, j'ai décrit la variabilité observée dans les expériences langagières parmi les bébés bilingues et j'ai discuté la façon dont ces expériences langagières peuvent possiblement influencer la segmentation des mots. Dans le *Chapitre 4*, j'ai recruté des bébés de 8 et 10 mois de familles unilingues et bilingues. Nos résultats confirment que

les bébés unilingues peuvent segmenter des mots à deux syllabes dans leur langue maternelle, mais pas dans une langue n'étant pas leur langue maternelle. Crucialement, nos résultats révèlent que les bébés bilingues peuvent, dès l'âge de 8 mois, segmenter des mots à deux syllabes dans chacune de leurs langues maternelles. Il est intéressant de noter que des analyses exploratoires suggèrent que dans notre tâche langagière bilingue, la capacité de segmentation de mots des bébés est renforcée si ceux-ci entendent plus de mélanges de langues par leurs parents.

Pour conclure, cette dissertation contribue à la littérature de plus en plus abondante mettant l'accent sur la grande variabilité en ce qui a trait aux expériences langagières bilingues et leurs effets sur les capacités de traitement précoce du langage. En effet, l'exploration des expériences langagières et habiletés des bébés bilingues nous offre une occasion unique d'étudier l'acquisition et le développement du langage.

ACKNOWLEDGMENTS

To my advisor Dr. Linda Polka, thank you for your constant support and guidance during my Ph.D. pursuit. Thank you for sharing your expertise and insights with me, and for providing me with a solid foundation on infant and language research. I truly appreciate and will always treasure your mentorship.

To the members of my various committees (Drs. Krista Byers-Heinlein, Elin Thordardottir, Aparna Nadig), thank you for taking the time to provide me with valuable feedback, and for strengthening the work in this dissertation. To Drs. Rachel Theodore and Kris Onishi, thank you for helping me grow as a researcher through collaborations and reading groups. To my undergrad advisor, Dr. Katherine White, thank you for getting me started on this rewarding journey, and for your mentorship, even today.

To the many families who participated in our research, thank you for spending your days with us, and for opening up your lives/homes to us. This research would not be possible without you.

To (past and current) members of the Infant Speech Perception Lab, thank you for all the discussions, and for all your hard work in supporting the research in this dissertation. Thank you especially to our lab co-ordinators (Alexina, Shila, Ying Ying, Paloma), and to our research assistants (Katherine, Julia, Jennifer, Jenn, Megan, Sabrina, Angela, Ellen, Melanie, and Vanessa).

To the trio (Dr. Colby: co-hort extraordinare; Dr. D: officemate extraordinaire; Alexandre: bad jokes extraordinare), thank you for being extraordinary at what you do. To my dissertation-writing cohort (Sarah, Amélie, Liz), thank you for the push to get it done, one page at a time. To my café buddies (Yonathan, Jamie), thank you for the good

company and distractions during the difficult writing tasks. Special mentions to my caffeine suppliers: Kafein, Anti-café, Café St-Henri and Café Aunja.

To my family (Dad, Mom, Ates, Kuya, Nephews) and friends, thank you for the daily encouragements, and for letting me escape work through Christmases, cottages, festivals, road trips and bike trips. To my soon-to-be wife, Johnalene, thank you for being there, every morneve and afternight. Your support means the world to me.

To God be all glory, honour and praise.

Acknowledgement of funding

PREFACE

The studies represented in Chapters 2, 3, and 4 are original contributions to knowledge. The first study (Chapter 2) is one of the first few studies to examine the use of the LENA recording system with a bilingual population, and the first to assess how the LENA algorithms deal with a bilingual speech stream. The second study (Chapter 3) is one of the first studies to examine the validity of parental reports in assessing a child's language environment, and provides new data describing the variability of bilingual input. The final study (Chapter 4) is one of the first few behavioral studies to examine early word segmentation in two languages, and the first to examine how word segmentation occurs in an inter-mixed dual language stream. These novel contributions provide us with a greater understanding of how bilingual infants acquire language from the input provided to them.

This dissertation is submitted as partial fulfillment for the Ph.D. program in Communication Sciences & Disorders at McGill University. It follows McGill University's guidelines for a manuscript-based thesis.

Contribution of Authors

Chapter One: Introduction

A.J. Orena solely authored this chapter, with feedback from L. Polka, K. Byers-Heinlein, and E. Thordardottir.

Chapter Two: Reliability of the Language Envionment Analysis (LENA) in French-English Bilingual Speech

This manuscript has been submitted for publication. It is co-authored by L. Polka and K. Byers-Heinlein. A.J. Orena developed the study concepts, designed the

experiment, collected the data, analyzed the data, and drafted the manuscript. L. Polka and K. Byers-Heinlein provided guidance during the development of the experiments, and provided feedback on the data analyses and draft of the manuscript.

Chapter Three: What do bilingual infants actually hear? Evaluating measures of caregiver speech to 10-month-olds

This manuscript has been submitted for publication. It is co-authored by L. Polka and K. Byers-Heinlein. A.J. Orena developed the study concepts, designed the experiment, collected the data, analyzed the data, and drafted the manuscript. L. Polka and K. Byers-Heinlein provided guidance during the development of the experiments, and provided feedback on the data analyses and draft of the manuscript.

Chapter Four: Segmenting words from inter-mixed bilingual speech during infancy

This manuscript has not yet been submitted for publication. It is co-authored by L. Polka. A.J. Orena developed the study concepts, designed the experiment, collected the data, analyzed the data, and drafted the manuscript. L. Polka provided guidance during the development of the experiments, and provided feedback on the data analyses and draft of the manuscript.

Chapter Five: General Discussion

A.J. Orena solely authored this chapter, with feedback from L. Polka, K. Byers-Heinlein, and E. Thordardottir.

CHAPTER 1. Introduction

In this dissertation, we examined the nature of bilingualism in infancy. This research topic has received growing attention in recent years, with many studies exploring how bilingual acquisition unfolds over the early formative years of life. The growth of research in this topic is, in part, in response to the growing bilingual population across the world, including in Canada (Statistics Canada, 2016). Indeed, much of our prior knowledge on language acquisition is based on infants growing up with one language. Thus, to obtain a more complete picture of language acquisition, many researchers are now examining how infants with other language backgrounds also acquire speech and language skills. This knowledge base is important not only for understanding language acquisition through a different lens, but also for informing research stakeholders (e.g., clinicians, educators, and parents) about the developmental trajectory and learning capabilities of young infants experiencing a common form of language acquisition (e.g., Byers-Heinlein & Lew-Williams, 2013; Core & Hoff, 2015).

Broadly, we investigated the relation between language input and speech processing skills in bilingual infants. Indeed, there is a myriad of research showing that infants who hear two languages from birth have a distinct path of speech and language development: for example, compared to monolingual infants, bilingual infants appear to have different speech perception and production patterns (Andruski, Casielles, & Nathan, 2013; Fennell, Byers-Heinlein, & Werker, 2007), and they use different heuristics for perceptual tasks, such as language discrimination (Bosch & Sebastian-Galles, 2003), speech sound categorization (Burns, Yoshida, Hill, & Werker, 2007) and word learning (Byers-Heinlein & Werker, 2009). These set of studies lay the important groundwork for

showing that the *bilingual experience* plays a formative role on speech perception and language acquisition. However, interpreting findings from these studies is often complex due to the wide heterogeneity in experiences among bilingual families. More research is needed to describe what the bilingual experience actually entails, and to understand how specific parameters of this experience shape infants' speech and language skills.

This dissertation aims to advance our knowledge by providing a more comprehensive story of the bilingual experience in infancy. To do so, we evaluated methods of assessing language input to bilingual infants and examined how input patterns relate to infant speech processing skills. We recruited both monolingual and bilingual families (N = 128 across three experiments) from Montréal, Québec, Canada to take part in our studies. This work is detailed in three manuscripts, which form chapters 2, 3, and 4 of this thesis. **Chapter 2** evaluates the utility of the Language ENvironment Analysis (LENA) recording system for examining bilingual speech in both a research and clinical context. Next, **Chapter 3** examines the use of the LENA recording system to assess measures of the dual language input of young bilingual infants, and to explore the heterogeneity of the bilingual experience in infancy. Finally, **Chapter 4** explores how this bilingual experience can affect infants' speech processing skills – specifically, their ability to recognize words in sentences (i.e., word segmentation). In **Chapter 5**, we integrate and summarize the findings across this body of work and discuss their implications for our understanding of bilingual acquisition. In the current chapter, we first briefly introduce the key themes that provide the motivation and context for this study including i) language input in infancy, and ii) early word segmentation.

1.1. Language input in infancy

As with many human processes, there is wide individual variability in language acquisition and growth: some children begin talking with a speedy learning trajectory, while others take a bit longer to get off the ground. For example, children differ greatly in *when* they start producing words and *how fast* they build their vocabulary (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). What is the source of these differences in language outcomes? Previous research points to a variety of predictors, including biological factors (e.g., Stromswold, 2001). For instance, a review of genetic studies posits that heritable factors play a strong role in the rate of language acquisition achieved by children (Stromswold, 2001). Nevertheless, even in the most extreme estimates, heritability of language development does not account for all of the variability in language outcomes; depending on the study's methodology, genetic factors account for only between 1 and 82% of the variability in language outcomes (Hoff, 2006). Thus, the environment must also play a role in how language develops in children. This dissertation focuses on one specific contributor from the environment: the *language input.*

Indeed, the input is important in any study of language acquisition. It is perhaps trivial to point out that some sort of speech and language input is needed for children to learn their native language(s). In the early stages of language acquisition, infants already make use of the structure of the input to adapt their speech-processing capacities for efficient language learning. For example, young infants already show phonetic discrimination patterns consistent with the phonological inventory of their native language (McMurray & Aslin, 2005; Werker & Polka, 1993), as well as a preference for phonemic, phonotactic, and prosodic features of their native language (Juszczyk, Cutler, & Redanz,

1993; Seidl, Cristià, Bernard, & Onishi, 2009). In fact, newborns already show language-specific patterns in phonetic perception (Zhao, Moon, Lagercrantz, & Kuhl, 2011), suggesting that language input during the pre-natal period already matters. Early vocalizations are also affected by the input: infants' cries and babblings tend to resemble the phonetic and prosodic characteristics of their native language (Levitt & Aydelott Utman, 1992; Wermke et al., 2017). These findings demonstrate that early changes in speech perception and production are not due solely to maturational changes in cognition and physical development; rather, these changes to infants' language processing are complex and involve the general language structure heard from the input.

In this section, we first describe how researchers solve the methodological challenge of assessing infants' language environment. Then, we describe how variability in language input affects infants' language development, with a focus on vocabulary development. As we describe in the following sections, we focused on the *input quantity* (in lieu of the *input quality*) given the current issues on examining the input quantity in bilingual infants, which we examine in **Chapters 2** and **3**; further, we focused on their effects on vocabulary development, given its relationship with word segmentation, which we examine in **Chapter 4**. We review the literature on these topics in the context of both monolingual and bilingual acquisition.

1.1.1. Assessing language input at home

Apart from extreme cases of deprivation, all infants in normal environments receive language input at home – whether it be from their parent or guardian, caregivers and/or other members of their household (Van de Weijer, 1999). Language researchers have used various methods – such as parental diaries, questionnaires, live observers, audio

and video recorders – to document how infants experience language. These different methods vary in terms of their practicality in data collection, as well as the depth of their measurements. The combination of these methods has provided researchers with a rich amount of data about how different families provide language input at home.

For example, interviewing parents about their child's language experiences can provide us with information about how different infants might hear different types of input at home. Depending on the research question, some researchers ask parents to fill out language diaries during a typical day, while others interview parents about their child's language environment during a laboratory visit. These reports show that infants' language environment can vary along many dimensions, including in the number of languages they hear at home (e.g., DeAnda, Bosch, Poulin-Dubois, Zesiger, & Frienda, 2016), the number of accents they are exposed to at home (van Heugten & Johnson, 2017), as well as the number of talkers that they frequently interact with (Bergmann & Cristia, 2018). Further, an infant's typical day can also vary in terms of how they are being cared for (daycare vs. caregivers at home) and the amount of time spent in different activities (e.g., book reading, media, rest; Place & Hoff, 2011). With respect to bilingual environments, researchers have also used parental reports to try to capture the extent to which bilingual parents mix their two languages to their infant (Byers-Heinlein, 2012). While these types of information are relatively easier to collect via parent reports, there can be some concerns about reporting bias. Indeed, there is limited research examining the validity and reliability of these parental reports.

One way to circumvent reporting bias is to directly observe infants' language environments via live or recorded observations. These observations can take place either

at the laboratory or at home, and they can be structured or more naturalistic. This method has given us additional information about the *quantity* and *quality* of language input that infants hear from their parents. For example, we have learned from these observational data that parents vary widely in the amount that they talk to their infants (input quantity; Hart & Risley, 1995), and that they also vary in the way they provide this input to their infants (input quality; Ramírez-Esparza, García-Sierra, & Kuhl, 2014). Prior studies have examined these characteristics of the language input in the context of different activities (e.g., book reading: Rvachew, Rees, Carolan, & Nadig, 2017), and how they might vary across different developmental populations (e.g., Down syndrome: Slonims & Mcconachie, 2006; Autism Spectrum Disorder: Wan et al., 2012).

While recorded observations provide more direct and accurate information about how infants experience language, they still carry some challenges. One limitation with these methods is that because parents are aware that they are being recorded, they may change their behaviors during these observation periods (e.g., talk more, "clean up" their speech). For example, one recent study showed that the observed language input is different when comparing speech samples collected via short video-recordings versus daylong audio-recordings (Bergelson, Amatuni, Dailey, Koorathota & Tor, 2018).

Another limitation is that they are logistically more difficult to conduct. Indeed, this method typically requires extra staff members to hold the recording equipment during the observation period. Further, depending on the research question, it can be time consuming to record, transcribe and code the data. To overcome these challenges, researchers have recently begun to use the **L**anguage **En**vironment **A**nalysis (LENA) recording and speech analysis system to collect data about infants' language

environments (LENA Research Foundation, 2014). This recording system from the LENA

research foundation allows researchers to record infants' auditory environment at home

for up to 16 hours. The system consists of a small child-friendly 2 oz. recorder that fits

into the front pocket of a t-shirt or vest that the child can wear all day. When the recording

is complete, it can be uploaded and processed through LENA software.

One feature of the LENA software is that it can automatically estimate several

characteristics of the recordings, including the number of words spoken by the parent and

child during the recordings, the amount of turn-taking between the parent and child, as

well as the amount of television time and other noisy sound sources. Further, it

automatically segments the recordings into speaker categories, environmental sound

categories, and silences using Gaussian mixture models. These types of information are

presented in the user-friendly software in both data and graph formats. The LENA system

has been used in English-monolingual homes to confirm the variability of the input

quantity to infants (Gilkerson & Richards, 2009). Further, even though the LENA

algorithms were designed for analyzing English recordings, its output for adult word

counts has been validated in other monolingual populations, including Spanish, European

French, Mandarin and Korean (Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2015;

Gilkerson et al., 2015; Pae et al., 2016; Weisleder & Fernald, 2013).

The LENA recording system was originally designed to be used in intervention

programs to promote effective interactions between parents and their children. As part of

their mission, the LENA foundation has introduced several programs, including LENA

Home (a home-visiting program), LENA start (a parent-group class), and LENA grow (for

early childhood educators), and they are also featured in several national initiatives in the

USA (e.g., Providence Talks: www.providencetalks.org/; The Thirty Million Words Initiative: https://tmwcenter.uchicago.edu/). Today, the LENA recording system is also used by language researchers and speech-language pathologists for examining the language input to young children.

For example, some researchers have used the LENA system to quantify the speech heard by different populations (e.g., preterm infants: Caskey, Stephens, Tucker, & Vohr, 2014; older adults: Li, Vikani, Harris, & Lin, 2014; children with Autism Spectrum Disorder Oller et al., 2010; children with Down Syndrome: Thiemann-Bourque, Warren, Brady, Gilkerson, & Richards, 2014). Other researchers are also examining how this variability in the language input affects different aspects of language development or speech perception (Vocabulary development: Caskey et al., 2014; GIlkerson & Richards, 2008; Speech processing efficiency: Weisleder & Fernald, 2013). A small handful of studies have also examined similar effects in a bilingual population (Marchman, Martinez, Hurtado, Grüter & Fernald, 2016; Garcia-Sierra, Ramírez-Esparza & Kuhl, 2016). Indeed, the LENA system provides us with a unique opportunity to investigate how different aspects of language input in the natural world affect infants' language outcomes.

Below, we briefly summarize the literature on how language input influences early language development. The characteristics of language input are often discussed in terms of its *quantitative* and *qualitative* properties; and both properties have been shown to be critical for various aspects of language development, including lexical development (e.g., Hoff & Naigles, 2002), and syntactic development (e.g., Thordardottir, 2015), as well as their cognitive and social development (Rowe, Leech, & Cabrera, 2016). For the

scope of this dissertation, we will focus on the effects of the input quantity on vocabulary development.

1.1.2. The effects of language input on vocabulary development

Monolingual acquisition

As described above, there is large variability in how parents provide language input to their children. This idea is perhaps most famously popularized by Hart and Risley (1995), who showed that a family's socioeconomic status (SES) played a role in the child's language development. In their seminal study, Hart and Risley collected hourly samples of language input in 42 American homes every month, from when the child was 7- to 36-months of age. They grouped their families into a *Professional* group (n=13), a *Working-class* group (n=23), and a *Welfare group* (n=6). Strikingly, Hart and Risley found a gradient relationship between SES and vocabulary score: children from a family with higher SES had larger vocabulary sizes. Hart and Risley hypothesized that one source of this relationship is that parents from these different SES groups provided different amounts of language input to their children.

Indeed, Hart and Risley's (1995) dataset revealed a positive correlation between SES and language input quantity: children from the *Professional* group heard, on average, 2,153 words per hour, while those from the *Welfare* group heard only 616 words per hour. They extrapolated the data to show that, by the time children enter kindergarten, some children would have heard almost 30-million more words than their peers – a phenomenon now known as the *30-million word gap* (Hart & Risley, 2003). This variability is quite striking, and clearly suggests a potential impact on infants' language outcomes. Conceivably, infants who hear more speech from their parents have significantly more

opportunities to learn the structure of their native language, and thus, have more opportunities to learn more words.

Indeed, this word gap appears to have significant effects on vocabulary development. Hart and Risley's (1995) findings revealed that by 3 years of age, children who received more language input from their parents had increased vocabulary size. Further, their follow-up study found that children's academic successes at age 10 were predicted by the amount of language input that they heard during their first 3 years of life (Hart & Risley, 1999). This research is supported by other work, which showed that infants who received more language input not only started to produce words earlier, but also had faster vocabulary growth (Hoff-Ginsberg, 1991; Huttenlocher et al., 1991). A recent report by Gilkerson and Richards (2009) showed that infants who received more language input from their parents scored consistently higher on standardized language assessments, such as the Child Development Inventory (CDI), Receptive-Expressive Emergent Language Test (REEL), and the Preschool Language Scale (PLS). Further, these children produced more daily vocalizations compared to other children who received less language input.

What is the rationale behind the notion that infants who receive more language input have better language outcomes? A data-providing view of the input would argue that infants who receive more language input have more opportunities to learn about the structures of their native language – which, in turn, promotes earlier lexical acquisition and faster vocabulary growth. As infants gain more exposure to their native language, they become more sensitive to the distributional regularities of their language (e.g., prosodic, phonemic, phonotactic properties) – which in turn, may allow them to extract

and learn about words in their language more efficiently. Indeed, prior work has shown that infants are excellent trackers of distributional information in speech (Maye, Werker, & Gerken, 2002), and that they are able to use this information for segmenting words in the speech stream (e.g., Jusczyk et al., 1993). Thus, it would follow that a higher amount of exposure would lead to more robust representation of language properties. Take the probabilistic cues for segmentation as an example: for transitional probabilities to be useful for segmenting words in natural speech, one would need a vast amount of input data to discover the syllable sequences with high transitional probabilities, that are found within words, and those with low transitional probabilities, that are found at word boundaries.

Certainly, frequency of the input affects lexical acquisition: infants' first words are typically those that they hear most often from their mothers (Harris, Barrett, Jones, & Brookes, 1988), and the order in which words appear in their lexicon is predicted by their input frequencies (Huttenlocher et al., 1991; Naigles & Hoff-Ginsberg, 1998). This finding is complemented by word learning experiments in laboratory settings (Swingley, 2007), as well as corpus analyses (Goodman, Dale, & Li, 2008). Further, several studies have shown that slightly older toddlers are faster to learn the words that they hear the most: correlational studies have shown that the frequency of exposure to a particular word strongly predicts learning of that word (Hoff & Naigles, 2002; Tamis-LeMonda & Bornstein, 2002). In a word learning task, Pinkham and colleagues (2011) showed that the majority of 4-year-olds (80%) could successfully learn an unfamiliar word after 24 repetitions of that word; but when the word was repeated only 3 times, only 20% of the children learned that word. Certainly, not all words have the same ease of learning –

nevertheless, this finding shows that children need repeated exposure of words to gain vocabulary, and that his could be accomplished by increasing overall input quantity to the child.

Another view of the role of input quantity focuses on the socio-pragmatic view of the input, which holds that word learning is improved when the parent and infant are mutually engaged in conversation. While the quantitative aspects of the input alone do not necessarily equate to high quality of parent-infant engagement, larger amounts of overall language input might be tightly linked to the amount of effort parents make to engage with their infant via speech. There is some empirical evidence to support this notion: Tamis-LeMonda and colleagues (1996) showed that infants who receive more verbal responses from their mothers show earlier vocabulary acquisition in both comprehension and production, and Akhtar and colleagues (1991) suggest that it is the frequency with which mothers respond to their infants that predicts vocabulary scores. While the frequency of language input is implicated in these results, it comes with a caveat – that is, the input quantity effects are stronger if the quality of speech is rich and engaging (e.g., Donovan, Leavitt, Taylor, & Broder, 2007; Ramírez-Esparza et al., 2014). Indeed, many studies have shown that various qualitative characteristics of the input matter for language acquisition. For example, recent findings have shown that it is the number of conversational turns between parent and child, and not the number of words, that is a better predictor of later language outcomes (Dwyer, 2017; Romeo et al., 2018). Thus, while we do not describe the effects of the input quality in detail here, it is important to note that the quality of the input is also important for many aspects of language learning.

Bilingual acquisition

Much of the research on input quantity discussed thus far has studied infants from monolingual homes. However, many children around the world learn two languages simultaneously from birth – an experience often referred to as *Bilingual First Language Acquisition* (BFLA; De Houwer, 2009) or *Dual Language Learning* (DLL; Genesee & Nicoladis, 2007). Historically, and even today to some extent, several concerns exist about raising a child with two languages simultaneously. From a developmental perspective, some researchers in the past worried that having to learn two languages at the same time may be too taxing for a child's language learning capabilities, leading to "confusion" of the two languages, or delays and impairments in language development (Leopold, 1949). Indeed, some educators and clinicians counsel against speaking to children with developmental disorders with more than one language (for further discussion, see Kay-Raining Bird, Lamond, & Holden, 2012; Paradis, 2007). Nevertheless, many studies today have confirmed that infants have the capacity to successfully learn two languages simultaneously (De Houwer, 2007). In fact, some studies show that monolingual- and bilingual-learning infants follow the same developmental trajectory in their early attunement to their ambient language(s) (e.g., Sundara, Polka & Molnar, 2008; Bijeljac-Babic, Höhle & Nazzi, 2016). These findings indicate that, in some contexts, infants can readily adapt their perception to more than one language. Further, studies have shown that being bilingual may bring cognitive benefits in executive functioning (Bialystok, Craik, Klein, & Viswanathan, 2004; but see Paap & Greenberg, 2013), which may be beneficial for children with developmental disorders (Gonzalez-Barrero & Nadig, 2017), as well as for protecting against the development of dementia later in life (Bialystok, 2011).

While the general principles of the input effects from monolingual research can be extended to bilingual-learning infants, it is well known that the nature of language input can differ widely between the two groups. Indeed, the fact that bilingual infants hear multiple languages at home leads to some distinct challenges for language acquisition. For instance, a unique aspect of bilingual speech is its quantity, since time spent on language acquisition is necessarily distributed between the two languages. Researchers typically measure the input quantity to bilingual infants by looking at the proportion of time the infant is reported to hear one language over the other. Studies using this measure reliably demonstrate that bilingual infants who received proportionally more exposure to a language had higher proficiency levels in that language. For example, Pearson and colleagues (1997) tested a group of bilingual infants aged 8 – 30 months and found that infants' percentage exposure to each language was a reliable predictor of their vocabulary development. This general finding is robust, and has been replicated with several bilingual populations, including Spanish-English bilingual infants (Place & Hoff, 2011), as well as French-English bilingual children (David & Wei, 2008; Thordardottir, 2011).

In general, considered within the same SES level, bilingual infants typically receive less input in each of their languages compared to a monolingual's native language (Garcia-Sierra, Ramírez-Esparza, & Kuhl, 2016). Does this affect bilinguals' overall vocabulary development? Several group comparisons have shown that simultaneously bilingual children have a developmental course similar to monolingual children, *but* only when the vocabulary inventories of both of their languages are considered (Patterson, 1998). Indeed, several studies show that bilinguals tend to fall behind in respect to monolingual norms when each language is considered separately (Hoff et al., 2012).

However, Thordardottir (2011) showed that this depends on the vocabulary measures used and the specific form of bilingual exposure. For example, their sample of 5-year-old simultaneously balanced (40 – 60%) bilinguals of English and French had comparable performance with monolingual children in receptive vocabulary; however, when expressive vocabulary was tested, these same bilinguals feel below the monolingual norms (but see Bialystok et al., 2010). In contrast, their dominant bilinguals (>60% exposure to one language) matched monolingual controls in expressive vocabulary in their dominant language, but not in their weaker language. Nevertheless, Thordardottir (2011) argues that although bilingual-learning infants take longer to reach monolingual norms, they do eventually catch up to them if there is sufficient language exposure.

However, as Marchman and colleagues (2016) noted, these proportional input measures do not paint a complete picture of the individual differences in language exposure between bilingual infants. There are at least two possible issues. First, just like among monolingual families, there is wide variation in the absolute amount that bilingual caregivers talk to their infants (De Houwer, 2011; Weisleder & Fernald, 2013). It is not yet clear in the literature how the relative proportion of exposure to each language aligns with the absolute amount of language input in each language, and how these two sources of variation in language exposure interact with proficiency levels in both languages. One could imagine that a balanced bilingual who hears 5,000 words per day in each language might have a different trajectory of language development compared to another balanced bilingual infant who hears only 500 words per day in each language. Indeed, early findings from two studies suggest that variation in language abilities of bilingual children can be better explained by the absolute amount of input in each language rather than the relative

proportion of exposure in each language (De Houwer, 2011; Marchman et al., 2016). These studies suggest that we should consider both types of input measures when examining individual differences in bilingual development.

Another issue is that proportion measures are often collected either via parent interviews or diary records (e.g., Bosch & Sebastián-Gallés, 2001), and as alluded to in an earlier section, little has been done to validate these parent-report measures. If we are to continue to use these parental reports in bilingualism research, it is necessary to ensure that parental reports do reflect the actual day-to-day experiences of bilingual infants. To our knowledge, only one study has investigated this issue. Marchman and colleagues (2016) asked Spanish-English bilingual parents how much of each language they spoke to their toddlers, and then collected day-long recordings of their language input to their toddlers using the LENA system. They found that the parental-reported proportion of exposure to each language was only marginally correlated with infants' raw amount of exposure to each language [$r = .46$, $p < .06$]. In any case, further efforts to improve input assessment methods are needed to build a better understanding of how input impacts bilingual development.

1.2. Early word segmentation

In the previous section, we described how the language input relates to vocabulary size for both monolingual and bilingual children. From a processing perspective, building this vocabulary entails many challenges for the language-learning infant, including the *word segmentation* problem (i.e., the ability to recognize chunks of the auditory stream as word forms). Indeed, as part of learning a novel word, one must be able to recognize this word whether in isolation or in a sentential context. Many researchers have argued

that this ability is an early pre-requisite for building a lexicon, with some studies showing

linking performance in a laboratory word segmentation task with both concurrent and later

vocabulary development (Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006;

Singh, Reznick, & Xuehua, 2012). The language-specificity of this ability motivates our

study for examining word form segmentation in bilingual infants in *Chapter 5*. In what

follows, we provide a basic overview of research on word segmentation in both

monolingual and bilingual infants.

1.2.1. Word segmentation in monolingual infants

Previous research shows evidence that young infants can recognize frequent and

highly meaningful words when presented in isolation: for example, infants as young as 4-

months of age can recognize their own names (Mandel, Jusczyk, & Pisoni, 1995).

However, words are rarely spoken in isolation (Aslin, Woodward, LaMendola, & Bever,

1996; Van de Weijer, 1999), and indeed, processing word forms (i.e., words to which the

infant may not have an attached meaning) in a sentential context is often thought to be a

more challenging task. The speech signal is a continuous auditory stimulus: unlike words

in written language, words in spoken language are not conveniently separated by spaces.

Even more challenging, acoustic breaks in speech do not necessarily coincide with word

boundaries (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Further,

processing a complete sentence involves more sophisticated lexical and syntactic

processing.

This challenging task of word segmentation has been the subject of numerous

studies. The early work by Jusczyk and colleagues showed that English-monolingual

infants are able to segment monosyllabic words between 6 and 7.5 months of age

(Jusczyk & Aslin, 1995), and bisyllabic words by 7.5 months of age (Jusczyk, Houston, & Newsome, 1999). Subsequent studies have also shown segmentation skills to emerge between 6 and 12 months of age in monolingual infants learning other languages, for example French (Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006), Dutch (Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000), and German (Hohle & Weissenborn, 2003).

So, how do infants, who have limited word knowledge, begin to tackle this great challenge of word segmentation? A large body of research suggests that infants tend to segment words from the bottom up, using a wide array of cues in the speech signal such as prosody, allophones, phonotactic patterns, as well as patterns of transitional probabilities, which we briefly describe below:

(i)   *Prosodic cues:* Studies show that infants are sensitive to the metrical structure of their native language (Jusczyk et al., 1993), and there is some evidence to suggest that this sensitivity begins with the infant's prenatal speech input (Nazzi, Bertoncini, & Mehler, 1998). It has thus been hypothesized that, like adults, infants can track the prosodic pattern of native language words. For example, given that the trochaic pattern (i.e., STRONG-weak words such as *DOC-tor*), is the dominant rhythmic pattern of content words in English, infants may find it easier to extract words that follow this pattern in a stream of speech. Indeed, word segmentation studies show that infants – as young as 7-months of age – can segment words that follow the predominant rhythmic pattern of their native language (Curtin, Mintz, & Christiansen, 2005; Jusczyk, Houston, et al., 1999; Linda Polka & Sundara, 2012).

(ii)   *Allophonic cues:* An early study by Hohne & Jusczyk (Hohne & Jusczyk, 1994)

showed that 2-month-old infants are able to discriminate between allophonic

variants of the same phoneme in their native language. A classic allophonic

example is the distinction between the /t/ in the pair "ni**t**rate" versus "nigh**t** rate": in

the former, when the /t/ is word-medial, it is aspirated and in the later, when the /t/

is word final, it is unaspirated. Thus, although both speech sounds represent the

same phoneme, they are realized differently depending on their position in a word.

As such, the acoustic features of speech sounds may be able to signify whether

they are at the onset or middle of a word. Indeed, there is some evidence to

suggest that 10.5-month-old infants can use this structural information to track the

boundaries of a word with such speech sounds (Jusczyk, Hohne, & Bauman,

1999).

(iii)   *Phonotactic cues:* Other studies show that infants are also sensitive to the

permissible (and disallowed) combination of phonemes in a familiar language

(Seidl et al., 2009). For example, sound sequences such as [mr] and [db] are

common as the syllable onsets in Polish, but never occur in English. This

phonotactic information may certainly be useful for word segmentation, given that

the aforementioned phonotactic sequences could only occur across word

boundaries in English. Indeed, 9-month-old infants are sensitive to how

phonotactic constraints can correlate with word boundaries (Mattys, Jusczyk,

Luce, & Morgan, 1999).

(iv)   *Probabilistic cues:* Infants can also use patterns of transitional probabilities to

support their word segmentation abilities (Graf Estes, Gluck, & Bastos, 2014;

Saffran, Aslin, & Newport, 1996). The rationale is simple: given that sounds in real bisyllabic words (e.g., *gui-tar*) co-occur more often than sounds that cross the word boundary (e.g. *tar-is* in the phrase "*the guitar is*"), it would follow that paired syllables with lower transitional probabilities signal the boundaries of a word, while paired syllables with higher transitional probabilities signal the presence of a word. Using non-word stimuli, some studies have shown that infants can indeed track distributional information (e.g., that two syllables always co-occur with each other) to determine word boundaries in speech (Saffran et al., 1996).

These studies suggest that infants can extract different types of information (i.e., prosodic, allophonic, phonotactic and probabilistic information) from the input to aid them in early lexical acquisition. Indeed, increased performance in these segmentation tasks has been associated with better concurrent word-learning abilities (Graf Estes, Evans, Alibali, & Saffran, 2007), as well as more advanced vocabulary later in development (R. Newman et al., 2006; Singh et al., 2012).

1.2.2. Word segmentation in bilingual infants

The language-specificity of word segmentation raises questions about how bilingual infants learn to segment words in each of their languages. Indeed, compared to monolingual infants, bilingual infants face several added challenges in terms of word segmentation, including receiving less input in each of their languages than monolinguals in their single language and having to learn how lexical forms vary in two systems instead of one. Recent computational studies suggest that segmentation performance from a bilingual corpus is only slightly disadvantaged, when compared to that from a monolingual corpus (Fibla Reixachs, 2017). Further, behavioral studies using artificial languages

suggest that adults and bilingual infants have the capacity to segment different words from two language streams with different properties (Antovich & Graf Estes, 2017; Weiss, Gerfen, & Mitchel, 2009). In fact, bilingual infants show a slight advantage over monolingual infants in these artificial language tasks (Antovich & Graf Estes, 2017; Tsui, Erickson, Thiessen & Fennell, 2017). However, very few studies have tested the segmentation skills of bilingual infants in natural languages.

Two findings in the literature provide evidence that bilingual infants follow the same developmental trajectory as monolingual infants with respect to segmenting monosyllabic words. Bosch and colleagues (2013) tested young monolingual infants and Spanish-Catalan bilingual infants in a word segmentation task. They found that, like their monolingual peers in their native language, both 6- and 8-month-olds bilingual infants were able to segment monosyllabic words in their dominant language. Another study by Singh and Foong (2012) showed that young English-Mandarin infants were able to segment monosyllabic words in both of their native languages by 7.5 months of age.

A recent study by Polka and colleagues (2017) examined bisyllabic word segmentation with 8-month-old bilingual infants acquiring English and French. In this study, a new protocol for testing segmentation was developed so that two different languages (i.e., English and French) could be tested within the same infant. Using this protocol, they confirmed that 8-month-old English and French-monolingual infants could segment words in their native language, but not in a non-native language, regardless of which language was tested first. However, the bilingual infants only showed successful segmentation for French, and only when it was tested first. Nevertheless, a different group of bilingual infants were able to segment words in English in a test protocol that provided

longer exposure to the stimuli, suggesting that bilingual infants are able to segment words in both of their languages in the appropriate testing condition. Previous studies suggest that longer familiarization times may reflect different processing strategies (Nazzi, Mersad, Sundara, Iakimova & Polka, 2014); thus, while bilingual infants are able to show segmentation in both of their native languages in certain testing conditions, these findings indicate that the strategies they use may be different than those of monolingual infants. Overall, these findings show that segmenting words in both their native languages in the dual-language task poses a distinct challenge for bilingual 8-month-olds, and that further research is needed to probe their processing capabilities for word segmentation.

1.3. The Present Research

In this manuscript-based dissertation, we explored how bilingual caregivers talk to their infants, and how this might affect their infants' language acquisition patterns. Specifically, we characterized what the bilingual experience might look like in a select group of bilingual families, and investigated how the global measures of the bilingual input are related to infants' early speech processing skills. To this end, we collected recordings of French-English bilingual caregivers talking to their 10-month-old infants at home (*Chapters 2 and 3*), and assessed infants' word segmentation abilities (*Chapter 4*).

We recruited all infants for this dissertation from the city of Montréal, Québec, Canada. The bilingual population in Montréal is unlike most bilingual communities around the world in that the two dominant languages present in Montréal (Canadian French and Canadian English) co-exist in many cultural and social contexts. Indeed, the latest census (2016) reports that 55.1% of individuals living in Montréal have fluency in both French and English, with about 6.0% of individuals reporting both French and English as their

first languages. Due to the cultural and political landscape of Montréal, there is a distinctively high amount of interest in the topic of language from the public and the media – especially in regard to language use in businesses and public space, in education, and in social services (Leimgruber, 2017). Thus, conducting research in Montréal provides us with a unique lens in investigating how children grow up in a bilingual environment.

To conduct these experiments, we used a combination of methods. To examine the language input, we used several methods, including standard parent-report interviews and a new recording and analysis technology by the Language ENvironment Analysis (LENA) foundation. The LENA recording system allowed us to collect full-day recordings of the infants' auditory environment at home. The data in these recordings provided us with a more complete picture with respect to how bilingual parents talk to their infants in a natural setting. The second is a classical paradigm used for testing infants called Headturn Preference Procedure (HPP; Nelson et al., 1995), which measures infants' preference for a set of stimuli over another based on their reactions to these sets of stimuli. The HPP paradigm has been used to examine a variety of processing skills, including language and phonetic discrimination, sensitivity to prosodic and phonotactic patterns, and many more. For the purposes of this dissertation, we used the HPP to examine word segmentation in young bilingual infants.

The research here was conducted for three main reasons. First, this dissertation aimed to advance methods for investigating language input in infancy. In **Chapter 2**, we assessed and validated the use of the LENA recording system for a bilingual population. The LENA software uses algorithmic models to approximate several measures of the input, such as total adult word count; and while previous studies indicate that these word

count estimates are reliable for English and European French separately, it is unknown whether the word count estimates would still be reliable when the input includes both Canadian French and Canadian English. While the LENA system has previously been used with bilingual populations (Spanish-English bilinguals: Marchman et al., 2016; Wood, Diehm, & Callender, 2016), no study has compared how well LENA's algorithms work across different languages, nor examined how well LENA's algorithms would work for bilingual input. Assessing these tools will be helpful in the future for asking interesting research questions that may not be as easily answered with more traditional research methods. Indeed, this is an important goal not only for research purposes, but also for clinical purposes. The LENA system has gained traction for being used as a language sampling tool in intervention programs for clinical populations, including children with hearing loss (Charron et al., 2016) and children with Autism Spectrum Disorders (Dykstra et al., 2013). Thus, validating the LENA's adult word count output for bilingual speech would ensure that the LENA device can be useful for bilingual clinical populations.

Second, with this dissertation, we aimed to provide more direct and detailed quantitative information about bilingual infants' language exposure. For practical reasons, many studies examining bilingual acquisition to date use parental reports to capture the bilingual exposure – typically in the form of proportional exposure to each language. In *Chapter 3*, we report our assessment of these parental reports, and describe some of the heterogeneity of bilingual experiences that even our seemingly homogeneous group of bilingual infants from Montréal might experience. Indeed, the linguistic landscape within bilingual families can differ in many ways, including in the raw amount of exposure to each language, the amount of exposure to language mixing, and the amount of exposure

to infant-directed and other-directed speech. These data will be helpful for guiding future questions on bilingual acquisition.

Finally, we aimed to show how the bilingual exposure might affect infants' speech processing skills. In **Chapter 4**, we examined the word segmentation skills of 8- and 10-month-old infants from different language backgrounds. Indeed, there is limited behavioral data that can address how bilingual infants achieve this important feat in language development. However, this is an important question, especially given the links between segmentation performance and later language abilities outlined above (Newman, Rowe, & Bernstein Ratner, 2015). Here, we presented our findings on the developmental trajectory of bilingual infants' segmentation abilities, as well as exploratory analyses showing how different types of language experiences (exposure to language mixing) might affect word segmentation. We hope that these data will fruitfully add to the growing literature on how bilinguals acquire their native languages.

CHAPTER 2. Reliability of the Language Environment Analysis (LENA) in

French-English Bilingual Speech

2.1. Introduction

A key factor in shaping early language development is the nature of early speech to infants. Indeed, many studies have established that the quantity and quality of caregiver speech has a strong influence on various subdomains of language development, including vocabulary growth (e.g., Hart & Risley, 1995; Ramírez-Esparza, García-Sierra, & Kuhl, 2014). Thus, evaluating and improving techniques for analyzing caregiver speech is important for many professionals, including researchers, clinicians and educators. Before the advent of modern recording systems, much of the research on caregiver talk was conducted with observation sessions and diary reports, which were limited by construct validity, parental bias, and human resources. A recent technological development from the Language ENvironment Analysis (LENA) foundation has attempted to mitigate these concerns, and has thus been the focus of many recent research and intervention programs focused on caregiver talk to young infants – initially with monolingual English infants, but more recently with other various populations, including monolingual learners of other languages, infants with identified disorders, and infants from a range of socioeconomic backgrounds (see Wang et al., 2017 for a review). Here, we focused on another understudied but populous group: bilingual infants. Indeed, more and more children around the world are learning two languages from birth, but many aspects of bilingual language acquisition remain poorly understood. The purpose of this

study is to assess the utility of the LENA recording system for examining caregiver talk to bilingual-learning infants.

The LENA recording system is a language measurement and analysis tool that assesses the auditory environment of a young child. The child-friendly hardware consists of a small 3 oz. wearable digital recorder that fits in the front pocket of a t-shirt or vest. The portable digital language processor (DLP) allows for the recording to take place in the infant's natural environment (i.e., at home), and without the need for bulky or visually salient equipment or extra personnel to be present. The DLP can record up to 16 hours of audio, which is then uploaded to software that automatically analyzes and segments the audio file. The LENA recording system generates reports about a child's language environment based on patented algorithms in the software. The report quantifies several aspects of speech in the recording, including the number of words spoken by adults, the number of vocalizations by the child, the number of conversational turns between the child and adult speakers, and the amount of noise and TV sounds in the background. The recording system was originally designed to support clinical and educational programs by providing quick feedback to caregivers, with the ultimate goal of increasing talk between caregivers and their children. Indeed, preliminary reports show that involvement in a program that uses the LENA recording system has positive effects for a child's language development (e.g., Suskind et al., 2013; Gilkerson, Richards & Topping, 2017). It has also been used by researchers to examine caregiver language input and children vocalizations in a range of ages and population type (Wang et al., 2017).

In the short time span since its inception, the use of the LENA recording system in research has both confirmed important data and revealed new findings about language

acquisition. The central focus of recent research is quantifying speech heard by different populations (e.g., preterm infants: Caskey, Stephens, Tucker, & Vohr, 2014; children with Down's Syndrome: Thiemann-Bourque, Warren, Brady, Gilkerson, & Richards, 2014; children with Autism Spectrum Disorder: Oller et al., 2010; older adults: Li, Vikani, Harris, & Lin, 2014), and exploring how changes in the amount of caregiver speech might affect different aspects of language development or speech perception. Several studies have confirmed the seminal findings of Hart and Risley (1995), that showed a tight link between the quantity of caregiver speech and monolingual infant's vocabulary development (e.g., Caskey et al., 2014; Gilkerson & Richards, 2008). Other studies have discovered a link between aspects of caregiver speech and speech processing efficiency (Weisleder & Fernald, 2013), and with children's brain responses to speech sounds (Garcia-Sierra et al., 2011; Romeo et al., 2018). Relevant to the current study, a small handful of studies have found the same type of input effects in young Spanish-English bilingual children (Speech processing efficiency: Marchman, Martínez, Hurtado, Grüter, & Fernald, 2016; Brain responses to speech sounds: Garcia-Sierra, Ramírez-Esparza & Kuhl, 2016).

An important consideration for the utility of the LENA recording system is the accuracy of its generated reports. The LENA recording system was originally developed for use in English environments, as the speech analyzer is based on American English recordings (Xu, Yapanel & Gray, 2009). Indeed, the LENA-generated adult word counts for English input are strongly correlated with word counts generated by a human transcriber ($r = .92$, $p < .01$; Xu et al., 2009). Given the many children in the world who grow up in other types of language backgrounds, it is of great interest to understand how well the algorithms, particularly for word counts, perform in non-English language

environments. Indeed, word forms are indexed differently across different languages, and it is possible that the algorithms developed for counting words in English may not be generalizable to other languages. Nevertheless, several studies have shown that LENA algoirthms can reliably count words in other languages, including European French (Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2015), Spanish (Weisleder & Fernald, 2013), Mandarin (Gilkerson et al., 2015), Korean (Pae et al., 2016), and Dutch (Busch, Sangen, Vanpoucke, & Wieringen, 2017). These studies found that the LENA-generated adult word counts were not necessarily accurate; they tended to underestimate the amount of word counts in the recordings. Nevertheless, all of these studies report that the estimated adult word counts for these other languages are sufficiently reliable, with correlation coefficients between .64 (European French) and .84 (Spanish). These findings suggest that the underestimation of LENA-generated adult word counts is consistent enough across participants to be used as a tool to compare the quantity of input across different infants.

Even though the LENA estimates for adult word counts have been validated for different languages, generalizing these results to a bilingual environment is not necessarily straightforward. First, it is difficult to assess whether the LENA-generated estimates are more reliable for one language over the other in bilingual input, especially given that previous validation studies have only examined the reliability of different languages separately, and these studies varied in their validation methodologies (see Table 2.1), including sample size and transcription period selection (see Ganek & Eriks-Brophy, 2017 for a short review), as well as data analysis (Busch et al., 2017). As an example, consider the two languages represented in this study: English and French.

Although these two languages share many cognates, they also differ in many linguistic properties (phonology, prosodic forms, speech rate; e.g., Pellegrino, Coupé, & Marsico, 2011). There have been efforts to validate the LENA-generated estimates of word count for both English and French (Xu et al. 2009; Canault et al., 2015); however, it is possible that the LENA algorithms are reliable at counting words within each of these languages, but are not comparable across the two languages.

Secondly, there are aspects of a bilingual input that might weaken the reliability of the LENA estimates. Indeed, the bilingual environment is not the sum of two monolingual environments (Grosjean, 1989), and validating the LENA system for two languages separately does not necessarily mean it would be reliable for the two languages in a bilingual environment. While the difference in acoustics of two languages in a bilingual context can be as wide as when measured separately (Danielson, Seidl, Onishi, Alamian & Cristia, 2014), there can still be varying degrees of accented speech and code-mixing in bilingual environments (e.g. Byers-Heinlein, 2012). In the original LENA technical report, Xu and colleagues (2009) acknowledged that the performance of the LENA algorithms can be affected by many sources of variability, including speaker variations (i.e., speaking style, speaking rate, speaker accent, pitch). Given that these speaker variations are often present in bilingual environments, it is important to examine how the LENA system deals with these challenges.

In sum, the goal of this study was to validate LENA's algorithms in a bilingual environment. As part of a larger project investigating input effects on language skills in bilingual infants from Montréal, Canada, we have collected LENA recordings of the auditory environment of 10-month-old infants hearing French and English. Here, we

assessed the reliability of the LENA-generated adult word counts across different languages (English vs. French), gender-tagged voices (Female vs. Male), and accented speech (Accented vs. Non-accented speech). While there is evidence of age-related changes in infant-directed speech (Kalashinkova & Burnham, 2018), previous validation studies have not found a difference in the reliability of the word count estimates across different ages (e.g., Busch et al., 2017; Canault et al., 2015; Xu et al., 2012). Thus, we believe that examining the reliability of the LENA-generated estimates for adult word count at one age point (i.e., 10-month-old) woud be sufficient for generalizing across different ages of the child. In the discussion, we report our assessment of the utility and limitations of LENA for investigating input in bilingual homes.

Table 2.1. List of previous studies examining the reliability of Adult Word Count across different languages

	Year	Language	Age	N	Transcriptions	r for AWC
Busch	2017	Dutch	2-5 y	6	48 five-minute segments	.87, p < .001
Canault	2015	French	3-48 m	18	324 ten-minute segments	.64, p < .001
Gilkerson	2015	Mandarin	3-23 m	22	66 five-minute segments	.87, p < .001
Pae	2016	Korean	4-16 m	99	63 ten-minute segments	.72, p < .001
Weisleder	2013	Spanish	19-24 m	29	120 fife-minute segments	.80, p <.001
Xu	2009	English	2-36 m	70	420 ten-minute segments	.92, p < .01

2.2. Methods

2.2.1. Participants

The participants were 21 families with a 10-month-old infant (13 males, 8 females), who were recruited as part of a large-scale study examining the language input to

bilingual-learning infants. All of the families lived in the Greater Montréal area in Québec, Canada, where over half of the population have knowledge of both French and English (57.44%; Statistics Canada, 2016). Parents reported no auditory or developmental neurocognitive disorders for their infants.

Family Information

During their first visit to the lab, the mean age of the infants was 9 months 29 days (range = 9 m 15 d – 10 m 14 d). All families consisted of one father (age range 27 to 46 years, *M* = 36.24) and one mother (age range = 30 to 41 years, *M* = 34.85). Ten infants had one older sibling, and two infants had two older siblings; the other nine infants were first-born single children. Mothers had an average of 17.9 years of education (range = 11 – 23), while fathers had an average of 17.1 years of education (range = 14 – 22). Our sample of families was from a mid to high socioeconomic background, with an average Hollingshead score of 52.2 (range = 31 – 66 out of a possible score of 66). Most of the infants were cared for at home full time by one or both of the caregivers (n=16), while a few spent significant time at home with a nanny (n=2) or were enrolled in full-time daycare (n=3).

During the family's visit to the laboratory, we conducted a language background interview with the family to estimate their child's exposure to each language (Byers-Heinlein et al., submitted). Parents did a month-to-month breakdown of their child's exposure to each of their languages, followed by a lifetime estimate of proportion exposure to each language. Then, we calculated the average of these two values. Based on these parent-report measures, there were 12 infants in a French-dominant language environment (range = 58 - 75% exposure to French), and 9 infants in an English-dominant

language environment (range = 57 - 76% exposure to English). Four families reported a small amount of exposure to a third language: Arabic, Kannada, Portuguese, and Spanish (1 - 5% exposure to the third language).

All caregivers reported having knowledge of both French and English, but the families diverged in terms of the strategies they used for speaking to their children. Some families reported that both caregivers spoke both English and French to their infant (n=9), others reported that one caregiver spoke both English and French and the other caregiver spoke only one language to their infant (n=8), and some reported using the one-parent, one-language strategy for their child (n=4). Overall, the participants represented a range of bilingual family language configurations.

2.2.2. Procedure

The first appointment took place in the laboratory. During this first session, we explained the purpose and the procedure to the families, gave them the materials for the study, and conducted the language background interview with one (n = 17) or both (n = 4) of the caregivers. Families were given three digital language processors (DLPs) and three vests designed to hold the DLPs. With these materials, families were asked to complete three full days of recordings at home – two on weekdays, and one on the weekend. Most families were able to follow this schedule, except for one family that recorded one weekday and two weekend days, and one family that recorded three weekdays.

At the end of each day of recording, parents were also asked to complete a daily activity diary to indicate the day's general activities. When the families finished doing the recordings, we scheduled another appointment – either at their home (n=19) or at the

laboratory (n=2) – to pick up the materials, complete more language questionnaires, and conduct a final interview. In all but two cases, both caregivers were present for the final appointment. The majority of families completed the experiment within two weeks (*M* = 12.71 days, *SD* = 5.59, range = 5 – 26 days).

2.2.3. Data analysis overview

The project provided us with 1008 hours of audio recordings (21 families X 3 days X 16 hours). Given the human resources required and time constraints, we implemented a two-stage protocol for selecting periods of the recordings from each family to transcribe. These protocols were modeled after methods used in previous studies (Marchman et al., 2017; Ramírez-Esparza et al., 2014). Accordingly, our protocol was designed to select periods of the recordings that contained different languages (French vs. English vs. Mixed), different social interaction types (speech to infant vs. speech between caregivers), and different degrees of speech density (high amount of speech vs. low amount of speech). In doing so, we aimed to transcribe and analyze periods that were representative of different scenarios in the whole recording.

Data coding and recording selection

First, we constructed a coding scheme to extract information from the recording about who was speaking and in what language. The LENA output provides an estimate of total speech heard by the infant, but it does not differentiate between what languages are being spoken and to whom the speech is being directed. The coding scheme was inspired by the Infant Social Environment Coding of Sound Inventory (SECSI; Ramírez-Esparza et al., 2014). We divided the recordings into 30-second chunks via Audacity software. Relevant to this study, trained research assistants listened to each chunk and

tagged each chunk for speaker context (i.e., who is speaking and to whom: *Mother, Father, Sibling, Infant, Other*), and language context (i.e., what language was being spoken: *English, French, Mixed, Unknown*). For the language context, if more than one language was being spoken during the 30-second chunk, research assistants coded it as "Mixed"; and if it was not clear what language was being spoken, we tagged it as "Unknown". To streamline this process, we only coded chunks that contained speech per LENA measures; in other words, chunks that had zero word counts per LENA measures were not coded. Further, based on pilot analyses, we found that coding half of the recordings was sufficient for capturing the language breakdown of the recording. Thus, we decided to code *every other* 30-second chunk that contained speech. Seven research assistants completed this part of the project. Each research assistant was a simultaneously and highly proficient bilingual of French and English. Each completed a training file before coding recordings to be submitted for analyses. We assessed the reliability among coders in these training files, and found that they were reliable in tagging the speaker (Mean$_{reliability}$= 94.2 %; Range = 91.8 % - 96.4%), and the language (Mean$_{reliability}$= 92.4 %; Range = 88.1 % - 96.1%). As part of their training, each research assistant was assigned to jointly code with one other research assistant for one coding session in order to maintain consistency across coders.

After this time-intensive process, we organized and aggregated the data into five-minute samples, so that we could identify the total amount of speech in each of these samples (per LENA measures), divided by social interaction type (per human coders), and in what languages (per human coders). For each participant, we identified nine samples of 5-minute recordings that would be representative of different degrees of

speech density (approximately the top 1st percentile and 50th percentile of adult word counts), directed to different interlocutors (infant, non-infant), and directed in different languages (English, French, and mixed). In total, we selected 189 five-minute samples (945 minutes of recordings) from 21 families to be transcribed.

Data transcription

Three research assistants completed the transcriptions after undergoing extensive training. One research assistant did a second pass of all transcriptions to ensure accuracy and consistency across all transcription files. The five-minute samples were transcribed in CLAN (Computerized Language ANalysis) software using modified CHAT transcription format. The LENA-generated file was processed through CLAN, which provides a framework for transcription. The file was segmented by the LENA algorithms into utterance-length segments, and it contained general information about who was speaking in each segment (Male adult speaker vs. Female adult speaker vs. Infant, as identified by LENA algorithms), and how many words LENA estimated were present for that specific segment. The research assistants orthographically transcribed each segment. They also identified the language of each segment, and counted the number words uttered in each language. A word was counted as such if it contained at least one syllable. To maintain consistency with previous work, we borrowed the rules set forth by Canault and colleagues (2015) to count words in both English and French. Specifically, free morphemes (e.g., *English*: the, a, an; *French*: le, la, les), prepositions (e.g., *English*: in, to, on; *French*: à, de, par), and pronouns (e.g., *English*: I, he, she; *French*: je, il, elle) were counted as one word. In addition, the elided forms were counted as part of the word to which it was attached (e.g., *English*: I'll, can't; *French*: aujourd'hui, l'chien).

2.3. Results

2.3.1. Descriptive statistics

There was considerable variability in the amount of talk among participating families. Per LENA's algorithms, our group of infants heard an average of 15,651 words from adults per day (SD = 8,190), with a wide range of 1,644 to 49,022 words in a single day.

We transcribed 189 5-minute samples (21 participants X 9 samples), for a total of 945 minutes of transcribed samples. Based on LENA algorithms, there were 50,419 LENA-generated segments within these 5-minute samples, of which 31.3% were female adult speech, 14.7% were male adult speech, 14.1% were infant speech, 6.3% were sibling speech, 2.0% were electronic sounds, and the remaining 31.65% were either other noise sounds or silence. We also coded the activities during these samples, and found that the primary activity in these selected samples varied, including play time (n = 56), conversations between adults (n = 41), bath or dressing time (n = 35), meal time (n = 22), story time (n = 27), media time (n = 2), and other housekeeping activities (n = 6).

2.3.2. Reliability of LENA for Adult Word Count (AWC)

Similar to previous studies, we observed that the adult word count (AWC) measure by the LENA software was an underestimate of the transcribed word count (see Table 2.2). On average, the words counted by LENA were 85.3% of those counted by transcribers (SD = 39.3%). To statistically compare these two measures, we used a non-parametric test (Wilcoxon signed rank test), as the distribution of values was not presumed to be normally distributed. Results showed that these two measures are significantly different when comparing all samples [V = 3411, $p < .001$]. Table 2.2 shows

that these measures are also significantly different within specific types of recorded input

based on language (*English* and *French*) and LENA-tagged sex of the voice (*Female* and

Male). These measures were not significant within samples that contained *Mixed* speech

[V = 2328.5, p = .11]. Note, however, that there were not many mixed utterances within

samples that contained language mixing, with a mean LENA-generated AWC of 18.46

(SD = 40.23) and a mean Transcribed AWC of 19.50 (SD = 36.76).

Table 2.2. Descriptive statistics of LENA-generated and transcribed adult word counts (AWC). The Wilcoxon signed rank test was used to compare these two measures. ** and * asterisks represent significance at p < .001 and p< .05, respectively.

	Number of samples	LENA-generated AWC	Transcribed AWC	Wilcoxon test
		Mean (SD)	*Mean (SD)*	*V*
All regions	189	339.51 *(179.22)*	423.34 *(210.76)*	3411 **
English	176	155.94 *(168.44)*	192.30 *(189.65)*	3746 **
Female	162	124.64 *(147.75)*	153.78 *(165.37)*	3133.5 **
Male	102	71.12 *(116.01)*	85.33 *(121.15)*	1494.5 **
French	179	160.74 *(154.12)*	198.17 *(177.01)*	3040 **
Female	167	107.38 *(131.48)*	133.08 *(150.73)*	2982 **
Male	112	96.68 *(119.31)*	115.88 *(120.40)*	1185 **
Mixed	106	18.46 *(40.23)*	19.50 *(36.76)*	2328.5
Female	77	20.08 *(46.32)*	20.07 *(42.28)*	1544.5
Male	46	8.92 *(11.72)*	11.09 *(11.02)*	350.5 *

By-language analysis

To examine the relationship between LENA-generated AWC and transcribed

AWC, we conducted a repeated measures correlation. This test was chosen to account

for intra-individual differences across the different participants. This analysis revealed that these two measures are strongly correlated [r_{rm}(167) = .77, p < .001] (see Figure 1a), suggesting that the LENA-generated AWC is a reliable measure of the actual AWC in the recordings.

We then assessed the relationship between automated and manual word count measures when separating the data by language. For each sample, we identified the segments that our researchers tagged as either English-only, French-only, and language-mixed (containing both English and French words). For these segments, we compared the LENA-generated versus the transcribed AWC. Note that some samples did not include any English-only segments (n=10), French-only segments (n=13), language-mixed segments (n=83), and were excluded from the respective analyses. The repeated measures correlation analyses reveal a significant relationship between the two measures for English-only segments [r_{rm}(154) = .90, p < .001], French-only segments [r_{rm}(157) = .94, p < .001], and language-mixed segments [r_{rm}(84) = .97, p < .001] (see Figures 1b, 1c and 1d), suggesting that the LENA algorithms were equally reliable in all language contexts.

To examine whether this relationship between LENA-generated and transcribed word counts is stronger for one language over the other, we conducted a linear mixed model, with the transcribed word count as the dependent variable, and LENA-generated word count and language as fixed effects, and participant and sample type as random effects. Prior to running the model, the language variable was rescaled and centered around zero. As expected, there was a significant main effect of LENA-generated word counts [β = 1.08, t = 43.33, p < 0.001], indicating a strong relationship between transcribed

and LENA-generated word counts in both languages. There was no main effect of language [β = -11.41, t = -.98, p = .33], indicating no significant difference in transcribed word counts between the English and French segments. Critically, there was no significant interaction between the LENA-generated word count and language [β = .07, t = 1.26, p = .21], suggesting that the relationship between the transcribed and LENA-generated word counts was consistent across language contexts.
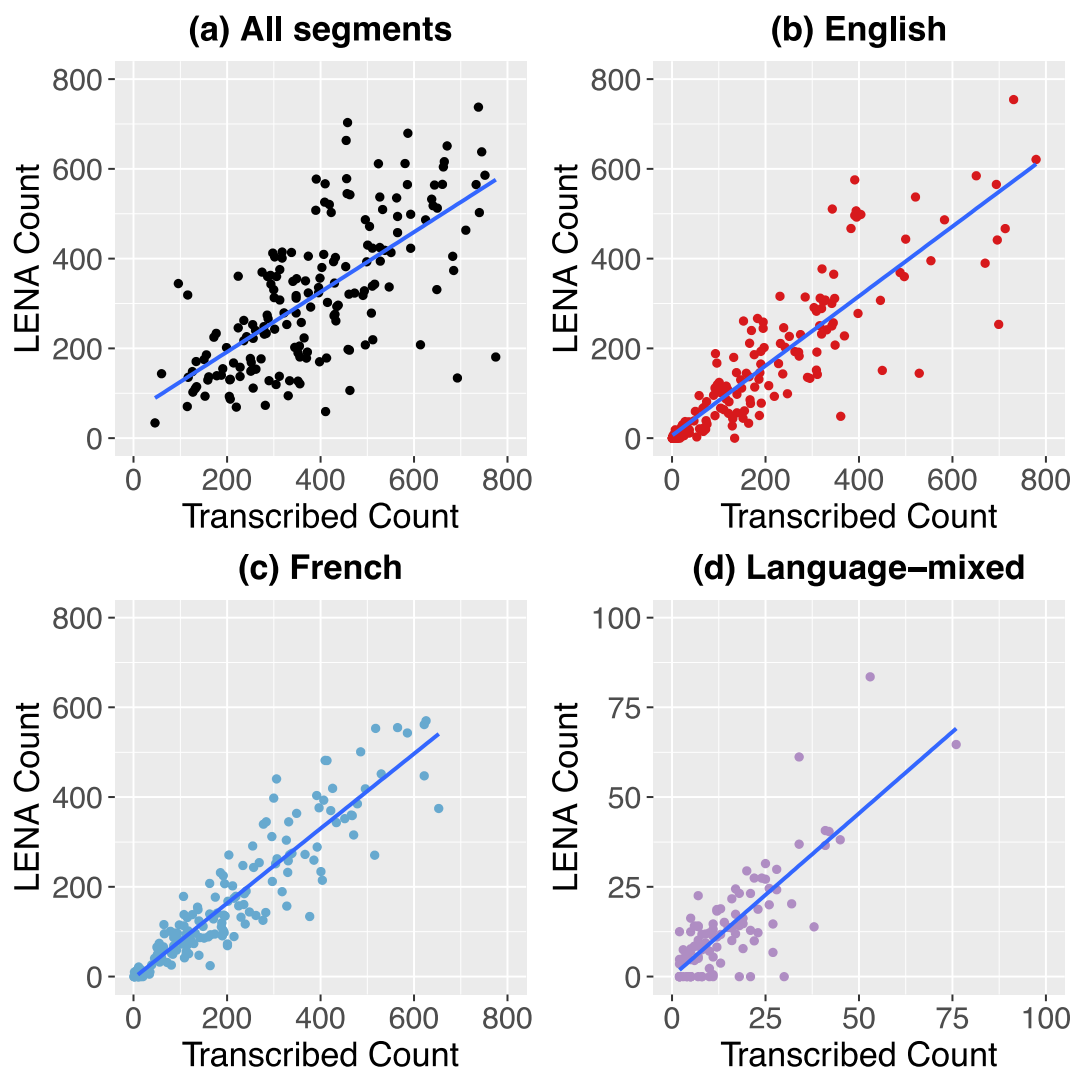

Figure 1. The relationship between LENA-generated and human-transcribed adult word counts for (a) all segments, (b) English-only segments, (c) French-only segments, and (d) Language-mixed segments. Each dot represents the word counts for one 5-minute sample.

By-gender analysis

We also examined whether the algorithms were similarly reliable for estimating word counts from female and male voices. To do so, we identified the segments that came from female and male voices, and then calculated an average for both the LENA-generated and transcribed AWC for each sample for each subject. Some samples did not include female (n=6) or male (n=55) voices, and were excluded from the respective analyses. Correlation analyses reveal a strong positive correlation between LENA-generated and transcribed AWC for both female voices [$r_{rm}(161) = .85$, $p < .001$] and male voices [$r_{rm}(112) = .90$, $p < .001$].

To examine this relationship more closely, we conducted another linear mixed model for this dataset, with the transcribed word count as the dependent variable, and LENA-generated word count and gender as fixed effects, and participant and sample type as random effects. Again, there was a significant main effect of LENA-generated word count [$\beta = .93$, $t = 28.39$, $p < 0.001$]. There was no main effect of gender [$\beta = -19.23$, $t = -1.22$, $p = .22$], nor an interaction between LENA-generated word count and gender [$\beta = -.07$, $t = -1.06$, $p = .29$], indicating that the relationship between the transcribed and LENA-generated word counts was consistent across the different-gendered voices.
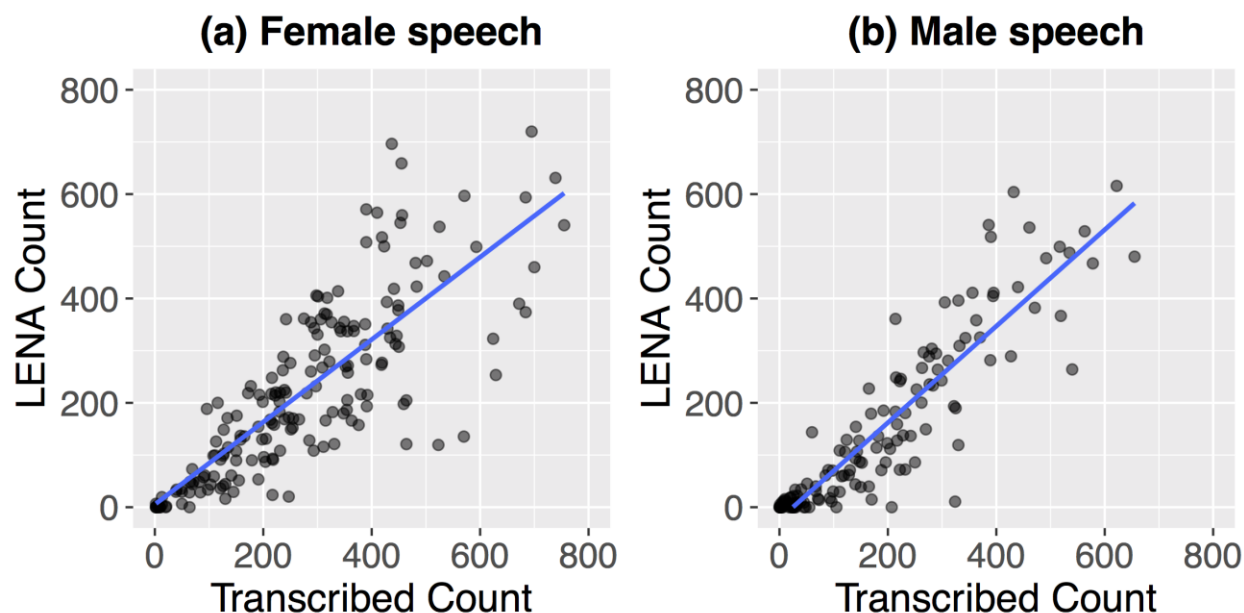
Figure 2. The relationship between LENA-generated and human-transcribed adult word counts for (a) segments with Female speech, and (b) segments with Male speech. Each dot represents the word counts for one 5-minute sample.

By-accent analysis

Finally, we examined whether the algorithms would be similarly reliable for estimating word counts from non-accented and accented speech. During the interview period, we asked parents whether they felt that they had an accent in either of their languages. For mothers, seven reported that they did not have an accent in either language, eight reported that they had an accent in English, and six reported that they had an accent in French. For fathers, three reported that they had no accent in either of their languages, fourteen reported that they had an accent in English, and four reported that they had an accent in French. Correlational analyses reveal a strong, significant relationship between LENA-generated AWC and transcribed AWC for both accented speech [$r_{rm}(217) = .88$, $p < .001$] and non-accented speech [$r_{rm}(375) = .86$, $p < .001$].

To examine whether the reliability of the LENA-generated AWC changes with accented speech, we conducted a linear mixed model, with the transcribed word count as the dependent variable, and LENA-generated word count and accent type as fixed effects, and participant and sample type as random effects. As expected, there was a significant main effect of LENA-generated word count [β = .96, t = 41.19, p < .001]. However, there was no main effect of accent type [β = 1.97, t = .35, p = .73], nor an interaction between these two factors [β = -.02, t = -.84, p = .40]. These data indicate that the LENA algorithms fare well for counting words even in accented speech.
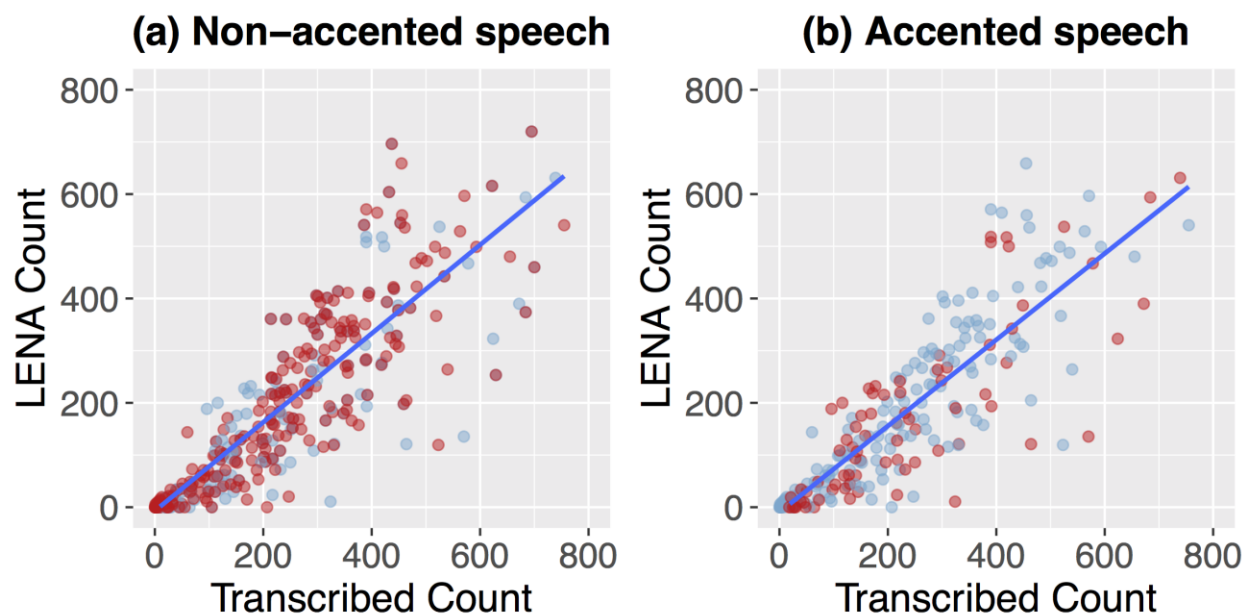

Figure 3. The relationship between LENA-generated and human-transcribed adult word counts for (a) segments with accented speech, and (b) segments with non-accented speech. Each dot represents the word counts for one 5-minute sample, with red dots representing English speech and blue dots representing French speech.

2.4. Discussion

In this study, we examined the utility of the LENA system for quantifying language input to bilingual infants. To do so, we assessed the accuracy of LENA algorithms for performing adult word counts in bilingual speech. The two languages featured here (Canadian English and Canadian French) differ in many phonological, prosodic and acoustic features, which raises the question of whether a system that is trained on American English phones can adapt to these features in a different language. Further, as is typical in bilingual settings, infants heard speech that contained both language mixing and accented speakers, which are speech and speaker features that may affect how well the LENA algorithms can estimate word counts in the recordings. Despite the potential challenges of measuring input in a bilingual context, we found high correlations between the LENA-generated and the human-transcribed adult word counts for both English and French – rs = .90 and .94, respectively, correlation coefficients are similar to those found by other researchers who examined LENA performance in different monolingual language environments. When we conducted separate analyses for different gendered voices and for accented speech, we also found a significant correlation between the LENA-generated and transcribed adult word counts. To our knowledge, this is the first study to directly compare the accuracy of the LENA system in a dual-language environment, and indeed, our results show that the algorithms in the LENA system adapt well to a bilingual speech stream.

It is important to note that the LENA system is more accurate for some recordings than others. In their original technical report, Xu and colleagues (2012) noted some systematic sources of error for the performance of the LENA algorithm, including the

presence of background noise, overlapping speech, clothing of the child, and so forth. Indeed, in our own exploratory analysis, we found that the overall correlations between LENA-generated and transcribed adult word counts were stronger for infants with siblings (r_{rm} = .88, p < .001) versus infants with siblings (r_{rm} = .69, p < .001). Importantly, results from the current study show that these systematic errors do not extend to different languages or accents, demonstrating the utility of the LENA system in a bilingual context.

Nevertheless, there are still some limitations and challenges of using the LENA system in a bilingual context. First, LENA currently has no way of automatically classifying utterances according to different languages. Thus, while it can provide information on input in aggregate, it cannot do so by language. As described in the methods, it was necessary for us to manually tag the language in the recordings, which was quite time-intensive and likely impractical for clinical purposes. For practical reasons, we recommend conducting language interviews with caregivers to estimate a child's proportional exposure to each language, and then combining this language exposure proportion with the LENA-generated measures to obtain a volumetric estimate per language. Indeed, preliminary results suggest that bilingual parents are sufficiently reliable at estimating their infant's language environment at home (Marchman, Martínez, Hurtado, Grüter, & Fernald, 2016; Orena, Byers-Heinlein & Polka, submitted). Adhering to instructions set forth by Byers-Heinlein and colleagues (submitted) may further decrease interview bias and increase the accuracy of the language exposure proportion.

Second, our study examined the accuracy of the LENA-generated word counts for Canadian English and Canadian French, and it is unclear whether our findings would generalize to the other LENA-generated estimates, as well as to other language pairs.

There are very few studies examining the reliability of LENA-generated frequencies of child vocalizations and turn-traking in other languages than English, but these studies indicate that these estimates are also not language-specific (e.g., Ganek et al., 2018; Gilkerson et al., 2015). Further, it would be interesting to examine whether the LENA algorithms are also reliable with other language pairs that differ in more linguistic properties than English and French (e.g., a pair of tonal and non-tonal languages). Future validation studies are necessary to fully capture the generalizability (or potentially, the lack of it) of the LENA algorithms to other language contexts.

Finally, the linguistic landscape of bilingual homes is very heterogeneous, and it may be challenging to capture a representative sample of this experience, even when using the LENA system. For example, bilingual children tend to receive their dual-language input from more than one speaker, which necessitates obtaining consent from more individuals to gather a representative sample. At the minimum, this study required the involvement of both parents, especially since some families reported using a version of the one-parent, one-language strategy. In addition, bilingual children tend to receive their dual-language input from other sources as well, including grand-parents, daycares, or community centers. Some families reported that they changed some of their activities during the day because they did not want to bring the recorders outdoors in public or because they did not want to have to involve their friends or family, which may have resulted in differences from the amount of each language the child might have heard in a more typical day. Further, issues related to obtaining consent prevented us from recruiting some families who had their child in full-day daycares. If researchers or clinicians are

interested in exploring the bilingual input as a whole, they may need to observe the child in a wider variety of contexts.

Nonetheless, our results indicate that using the LENA recording system in bilingual homes does yield a valid representation of the bilingual language landscape. At the end of the experiment, parents were asked a series of questions about their experience with the LENA recording device. When asked whether they felt that they changed the amount of speech they spoke to their child, only 14 out of 40 caregivers said "yes"[1]. Of these caregivers, ten said that they spoke to their child more often. For example, one parent noted that "*Once in a while, … if it was quiet, I'd be like "oh yeah, we're recording!" and I guess I should say something, so I would talk a tiny bit more.*" Four other participants said that they spoke to the child less often, indicating that they were a bit more cautious about discussions within the family. For example, one parent cited the *Hawthorne effect*, noting that *"there's something about… having a microphone that makes you more conscious about what you're doing".* Nevertheless, these parents reported that, while they acted differently at the beginning of the recordings, they quickly resumed their daily activities at home.

While we encouraged parents to act as natural as possible during the recordings, some parents did note that they changed the language proportion that they spoke to their child (n=7). For example, one parent said, *"It [the study] almost just serves to remind me, like, I need to speak more English, and I'm hoping it, like, continues from now on"*, and one other parent said, *"I probably changed by putting more English to it [the recording]…*

---

[1] Note that two caregivers were not present for the post-study interview; thus, we only report interview data for 40 out of 42 caregivers in this study.

It's more of a realizing that, oh, maybe we don't do it [speak English] as much." This

suggests that, in addition to increasing the quantity of language input at home, the tool

could also be used as a way to change the language proportion at home to be closer to

parents' goals. Indeed, several studies have shown that differential patterns of language

exposure affect various speech and language outcomes, including speech processing

skills (Hurtado et al., 2013) and vocabulary development (Thordardottir, 2011). Thus, it

may be beneficial for parents to know the raw amount of input that their child hears in

each language. Future studies should examine the potential utility of this tool for

improving bilingual language outcomes in children.

In sum, the LENA recording system offers researchers and practitioners a means

to investigate a child's language input at home. Our study shows that the use of the

system can be extended to French-English bilingual families. From a research

perspective, this tool can answer more detailed questions about how specific parameters

of the bilingual input affect language skills, which can inform theories of language

development (Odean, Nazareth, & Pruden, 2015). For example, the LENA system is able

to provide the absolute amount of input that a child hears, which has been shown to be

predictive of a bilingual child's speech and language skills (Marchman et al., 2016;

Garcia-Sierra et al., 2016). From a practical perspective, this tool can be used for

improving caregiver talk to bilingual infants. Indeed, many children grow up learning two

languages, so it is important to develop tools that can be used in both monolingual and

bilingual environments. Certainly, the system has been shown to be effective in increasing

caregiver talk, even in different cultures (Benítez-Barrera, Angley, & Tharpe, 2018; Pae

et al., 2016). On both ends, the LENA recording system opens up possibilities for investigating caregiver talk to children and improving their child's development.

Acknowledgments

2.5. References

Benítez-Barrera, C. R., Angley, G. P., & Tharpe, A. M. (2018). Remote Microphone System Use at Home: Impact on Caregiver Talk. *Journal of Speech, Language & Hearing Research*, *61*(2), 399–409. https://doi.org/10.1044/2017_JSLHR-H-17-0168

Busch, T., Sangen, A., Vanpoucke, F., & Wieringen, A. van. (2017). Correlation and agreement between Language ENvironment Analysis (lena™) and manual transcription for Dutch natural language recordings. *Behavior Research Methods*, 1–12. https://doi.org/10.3758/s13428-017-0960-0

Byers-Heinlein, K. (2012). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism*, *16*(1), 32–48. https://doi.org/10.1017/s1366728912000120

Byers-Heinlein, K., Schott, E., Gonzales-Barrero, A. M., Brouillard, M., Dubé, D., Jardak, A., … Tamayo, M. P. (submitted). Best practices for using parent report to measure bilingual infants' language environments.

Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2015). Reliability of the Language ENvironment Analysis system (LENA™) in European French. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-0634-8

Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2014). Adult Talk in the NICU With Preterm Infants and Developmental Outcomes. *Pediatrics*, *133*(3), e578–e584. https://doi.org/10.1542/peds.2013-0104

Danielson, D. K., Seidl, A., Onishi, K. H., Alamian, G., & Cristia, A. (2014). The acoustic properties of bilingual infant-directed speech. *The Journal of the Acoustical Society of America, 135,* EL95. https://doi.org/10.1121/1.4862881

Ganek, H. V, & Eriks-Brophy, A. (2017). A Concise Protocol for the Validation of Language ENvironment Analysis (LENA) Conversational Turn Counts in Vietnamese. *Communication Disorders Quarterly*. https://doi.org/10.1177/1525740117705094

Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2016.10.004

Gilkerson, J., & Richards, J. A. (2008). The LENA foundation natural language study (Technical Report LTR-02-2). Retrieved from: http://www.lenafoundation.org/TechReport.aspx/Natural_Language_Study/LTR-02-

2

Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., … Topping, K. (2015). Evaluating Language Environment Analysis system performance for Chinese: A pilot study in Shanghai. *Journal of Speech, Language, and Hearing Research*, *58*(2), 445–452. https://doi.org/10.1044/2015_JSLHR-L-14-0014

Grüter, T., Hurtado, N., Marchman, V. A., & Fernald, A. (2014). Language exposure and online processing efficiency in bilingual development. *Input and Experience in Bilingual Development*, *13*, 15. https://doi.org/10.1075/tilar.13.02gru

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H Brookes Publishing.

Kalashnikova, M. & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar properties but different functions. *Journal of Child Language, 45*(5), 1035-1053. https://doi.org/10.1017/S0305000917000629

Li, L., Vikani, A. R., Harris, G. C., & Lin, F. R. (2014). Feasibility study to quantify the auditory and social environment of older adults using a digital language processor. *Otology and Neurotology*, *35*(8), 1301–1305. https://doi.org/10.1097/MAO.0000000000000489

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2016). Caregiver talk to young Spanish-English bilinguals: comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*. https://doi.org/10.1111/desc.12425

Odean, R., Nazareth, A. & Pruden, S. M. (2015). Novel methodology to examine cognitive and experiential factors in language development: combining eye-tracking and LENA

technology. *Frontiers in Psychology, 6*(1266). https://doi/org/10.3389/fpsyg.2015.01266

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., … Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, *107*(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107

Orena, A. J., Byers-Heinlein, K., & Polka, L. (submitted). What do bilingual infants actually hear? Evaluating measures of caregiver speech to 10-month-olds.

Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J. A., Ma, L., & Topping, K. (2016). Effects of feedback on parent–child language with infants and toddlers in Korea. *First Language*, *36*(6), 549–569. https://doi.org/10.1177/0142723716649273

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language, 87*(3), 539-558. https://doi.org/10.2307/23011654

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

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2016). The Impact of Early Social Interactions on Later Language Development in Spanish-English Bilingual Infants. *Child Development*. https://doi.org/10.1111/cdev.12648

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

Statistics Canada (2016). Focus on Geography Series, 2016 Census. *Statistics Canada Catalogue no. 98-404-X2016001*. Ottawa, Ontario.

Thiemann-Bourque, K. S., Warren, S. F., Brady, N., Gilkerson, J., & Richards, J. A. (2014). Vocal interaction between children with down syndrome and their parents. *American Journal of Speech-Language Pathology*, *23*(3), 474–485. https://doi.org/10.1044/2014_AJSLP-12-0010

Wang, Y., Hartman, M., Aziz, N. A. A., Arora, S., Shi, L., & Tunison, E. (2017). A Systematic Review of the Use of LENA Technology. *American Annals of the Deaf*, *162*(3), 295–311. https://doi.org/10.1353/aad.2017.0028

Weisleder, A., & Fernald, A. (2013). Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENATM Language Environment Analysis System in young childrens natural home environment. *Behavior Research Methods* 48(3), 1109–1124. https://doi.org/10.3758/s13428-015-0634-8

Preface to Chapter 3

The study in Chapter 2 examined the utility of the LENA (**L**anguage **En**vironment **A**nalysis) recording system for investigating the language input to bilingual infants. To do so, we asked 21 bilingual families with a 10-month-old infant to record three full days of their lives using the LENA system. The results of Chapter 2 showed that the LENA system provides a reliable and valid estimate of word counts for both Canadian English and Canadian French, even in the presence of gender and accent variability. This finding allows us to be confident in using the LENA-generated values for other analyses.

In Chapter 3, we examined the LENA-generated input analyses of the bilingual infants from Chapter 2 in more detail. First, we assessed another methodology for examining the language input in infants: caregiver interviews. Specifically, we asked whether caregivers could reliably estimate their child's proportional exposure to each language, as observed in the LENA recordings. Second, we described the variability in how bilingual caregivers provided language input to their infants. These analyses represent the next step in investigating bilingual infants' language experiences.

CHAPTER 3. What do bilingual infants actually hear? Evaluating measures of caregiver speech to 10-month-olds.

3.1. Introduction

Children's speech and language outcomes depend, in part, on how their caregivers provide language input (Hart & Risley, 1995). Infants growing up in a bilingual environment encounter this input in different languages, in different amounts, and from different people (De Houwer, 2007). One of the methodological challenges for developmental researchers is capturing this "bilingual input" in a meaningful way. For example, many researchers conduct interviews with parents to estimate a bilingual child's proportion of exposure to each of their two languages. Using such measure, studies have found that different patterns of language exposure affect speech perception (Bijeljac-Babic, Serres, Höhle, & Nazzi, 2012), speech processing (Hurtado, Grüter, Marchman, & Fernald, 2014), and vocabulary development (Thordardottir, 2011). But, how valid are these parent reports? Are they a reliable representation of what the child actually experiences at home? And what aspects of the bilingual input are, or are not, captured by parent-report measures? The overarching goal of this paper is to investigate the nature of dual language input to bilingual infants, and to address some of the methodological concerns about parent reports in research on bilingual development. First, we will assess the accuracy of parent reports in measuring bilingual infants' language exposure. Second, using a data-centered approach, we will describe some of the heterogeneity of bilingual exposure in a sample of 10-month-old, French-English bilingual infants from Montréal, Québec, Canada.

A frequent and longstanding question in the study of bilingualism is how the pathway of bilingual development compares to monolingual development. The language environments of bilingual and monolingual infants are characteristically different from each other (see Byers-Heinlein & Fennell, 2014 for a review), resulting in differing language-learning contexts. Indeed, the presence of two languages in the environment presents bilingual infants with extra processing challenges, including having to learn and perceptually represent the linguistic properties of two language systems (Werker & Byers-Heinlein, 2008). In some cases, bilingual infants demonstrate language competencies comparable to their monolingual peers, including in phonetic categorization (Sundara, Polka, & Molnar, 2008), prosodic acquisition (Bijeljac-Babic, Höhle, & Nazzi, 2016), and word learning (Fennell & Byers-Heinlein, 2014; Mattock, Polka, Rvachew, & Krehm, 2010). In other cases, differences between these two groups are observed, such as in phonetic discrimination (Bosch & Sebastian-Galles, 2003), and in audiovisual perception (Pons, Bosch, & Lewkowicz, 2015).

These group-level comparisons between monolingual and bilingual infants lay the important groundwork for showing that dual language exposure can impact various aspects of language development. A different approach to investigating bilingual development is exploring how individual patterns of bilingual exposure affects language outcomes. Indeed, along with the inherent differences between monolingual and bilingual environments, there is also wide variability in dual language exposure *within* bilingual infants. For example, bilingual families can differ in the pair of languages that they speak, the caregiver's nativeness and fluency in each language, the amount of talk infants hear in each of their languages, and the contexts in which infants hear these languages.

Further, the language landscape across bilingual homes can differ quite vastly, with families adopting various language use patterns at home (e.g., one-parent, one-language structure; King, Fogle, & Logan-Terry, 2008). All these experiential factors may affect a child's speech and language outcomes, raising the importance of examining the effects of bilingualism beyond group-level comparisons (Luk & Bialystok, 2013; Titone & Baum, 2014).

One predictor variable that has received much attention is a bilingual child's quantitative experience with each of their languages. For example, prior research has shown that monolingual infants who hear more talk from their caregivers tend to have larger vocabularies (Hart & Risley, 1995; Weisleder & Fernald, 2013). One might expect then that bilingual infants would lag behind in their single-language development. As a group, bilingual children tend to have smaller vocabulary sizes and slower rates of vocabulary growth over time, when considering vocabulary size in only one of their two languages (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Place & Hoff, 2011; Vagh, Pan, & Mancilla-Martinez, 2009). Nevertheless, when taking both of their languages into account, bilingual children typically have the same rate of vocabulary growth as monolingual children (Core, Hoff, Rumiche, & Señor, 2013; Pearson, Fernández, Lewedeg, & Oller, 1997).

Less is known about how the relationship between bilingual exposure and language outcomes is reflected on an *individual level*. This is partly due to the logistic difficulties in capturing the absolute quantity of language input to bilingual infants (i.e., the number of words or utterances heard in each language). Some researchers have tried to measure this variable by observing parent-child interactions in a laboratory setting, or by

recording a child's auditory environment at home in a typical day (Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). However, the absolute quantity of language input is much harder to capture for bilingual children, since the bilingual input typically comes from more than one parent and occurs in different contexts. For example, a child might hear one language more often from a primary stay-at-home caregiver, and only hear the other language from a working caregiver during evenings and weekends. These contexts make it difficult for researchers to capture a child's bilingual exposure within a single observation period.

Due to the logistic difficulty of capturing the absolute quantity of child-directed speech in both languages, most researchers estimate a bilingual child's relative exposure to each language instead (i.e., proportion amount of time being exposed to Language A versus Language B). One practical way to assess this proportional variable is to elicit parent reports – either via diary reports (Place & Hoff, 2011), or detailed interviews with the parents (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Frienda, 2016; Byers-Heinlein et al., under review). Indeed, researchers have found a relation between some variant of this proportional measure and various speech and language outcome measures. For example, some researchers use these parent reports to divide their bilingual samples into two groups: "balanced" infants, who receive relatively equal input in each of their languages, and "dominant" infants, who receive relatively more input in one of their languages over their other (e.g., Bijeljac-Babic et al., 2012). Other researchers have also used a variant of this categorical measure (e.g., dominance in Language A vs. dominance in Language B) to examine differences in other speech processing tasks (e.g., Polka, Orena, Sundara, & Worrall, 2016; Sebastián-Gallés & Bosch, 2002). Finally, some studies

have used parent-report measures on a continuous scale to examine a potential dose-response relationship between language exposure and language outcomes. Indeed, the proportion of language input is predictive of a bilingual child's vocabulary development in each of their language at two years of age (Hoff et al., 2012), as well as their ability to efficiently process speech in that language at three years of age (Marchman, Martínez, Hurtado, Grüter, & Fernald, 2016).

One methodological limitation is that it is not clear how reliable these parent reports are in estimating a child's bilingual experience. It is possible that the reason we find relations between language exposure and language outcomes is that parents are biased by their child's proficiency in each language. That is, reports of time spent with each language might be colored by the parent's observation and interpretation of their child's communicative competence with them. Further, while it is common and standard practice in the health field to use parent reports as a proxy for child assessments (e.g., Theunissen et al., 1998), there are particular challenges that come with estimating a bilingual child's language experience. First, it may be difficult for the average bilingual speaker to track their on-line use of different languages (Carroll, 2015). In certain contexts, language mixing can occur quite often, and there is not a clear separation between languages. Thus, parents who mix their languages more often may not be able to provide as accurate information about their child's language environment as parents who use a single language. Second, estimating a child's language environment might be especially difficult when the parents are not around the child all the time, for example if the child attends daycare. Are parent reports of language exposure only representative of their own talk to their children, or do they also capture the language input spoken by other people? While

parents tend to provide the vast majority of the input to their child, several studies show that speech by others present in a child's environment (e.g., siblings, grandparents) also matters for their language outcomes (Bridges & Hoff, 2014)

Two recent studies have attempted to address these questions. As part of a longitudinal study, De Houwer and Bornstein (2016) asked Dutch-French bilingual mothers of young children to identify what languages they spoke to their child, and then video-recorded the mother and child interacting in both structured and non-structured sessions. They concluded that, for the most part, mothers were accurate at identifying what languages they spoke to their child. When their child was 5 months of age, mothers' reported language use matched what they spoke in a 60-minute non-structured interaction with their child (24 out of 24 mothers matched). Similarly, when their child was 53 months, their reported language use matched what they spoke in structured reading and play tasks (22 out of 24 matched). However, when their child was 20 months of age, their reported language use did not match what they spoke in a 20-minute meal and play session (only 12 out of 24 matched). For these mismatches, the error was due to mothers reporting that they spoke only one language to their child, but actually spoke both languages in the recorded session.

As part of a larger study on bilingual speech processing, Marchman and colleagues (2016) also examined whether parents are accurate at estimating how much bilingual children hear each language. Researchers first interviewed Spanish-English bilingual parents of young children (36- to 40-month-olds) about their child's language experiences (i.e., how many hours per day their child hears each language). Then, they made use of a recording system developed by the LENA (Language Environment

Analysis) foundation to examine the child's language environment at home. This recording system consists of a small recording device that fits into a vest or T-shirt that the child can wear over a full day, as well as software that can automatically estimate the total number of words being spoken to the child. Because the software cannot currently discriminate between the languages in the recording, the researchers listened to the recordings and tagged them for language. The results suggested a positive, albeit moderate, relationship between parent reports of language exposure and the observed proportion of exposure to each language in the recordings [rs(16) = .46, p <.06].

These two studies provide some support for using parent reports to examine bilingual children's language input. However, there are some limitations in regard to the generalizability of their findings. Firstly, De Houwer and Bornstein (2016) focused specifically on mothers' own language use with their children. However, children receive language input from multiple different speakers in the household (other caregivers, siblings, grandparents), and it is not clear from their study whether parents can also estimate what their child hears from these other speakers. Secondly, in both of these studies, the recordings were mostly conducted in one context (i.e., one weekday at home, or one session at the lab). However, as Marchman and colleagues (2016) discussed, these recordings or observations may not represent the child's range of language experiences, especially since children at these ages typically spend some time outside the home (e.g., at daycare, or at other family's homes), where language input might be very different. Thus, the observations or recordings may not have measured the same range of contexts that parents were able to capture in their reports, which may account for some of the mismatch. Finally, in both of these studies, parents tended to speak only

one language to their child. Thus, it is unclear whether parent reports of bilingual families who speak both languages to their child would also be similarly accurate. Indeed, it is possible that there may be greater discrepancy between reported and observed measures of language experience when there is more language mixing in the environment.

In this paper, we continue to tackle these issues. First, we assessed the validity of parental reports in a group of French-English bilingual families from Montréal, Québec. The bilingual population in Montréal is unique in that the two dominant languages (French and English) co-exist in many cultural and political contexts. Over half of the population in Montréal report being fluent in both French and English (55.1%; Statistics Canada, 2016), and language mixing is prevalent in many social contexts. Thus, testing this population will provide us with new insights on how linguistically-aware individuals can estimate their child's language environment. Second, we explored the nature of how bilingual parents spoke to their children, specifically in regard to: i) consistency of language use and ii) differences in language use between infant- and other-directed speech. There is substantial evidence that these differential patterns of language exposure affect language outcomes (e.g., De Houwer, 2007; Thordardottir, 2011). As such, it is important to examine the language input to bilingual infants with more depth, and consider current practices in assessing language input to bilingual infants.

3.2. Methods

3.2.1. Participants

Twenty-one families with a 10-month-old infant (13 boys, 8 girls; age range = 289–319 days, M = 303 days) took part in this study. This study was conducted as part of a

larger project examining the effects of language input on speech processing at 10-months-of-age. Validating parent reports at this age is useful, given that many speech processes develop at this age range (e.g., word segmentation: Jusczyk, Houston, & Newsome, 1999; attunement to native sounds Polka & Werker, 1994). We recruited our participants from a database of infants who were born in Montréal, Québec. We aimed to recruit infants who heard both French and English between 25% and 75% of the time, without exposure to another language more than 10% of the time. We explained the premise and basic procedures of the study to the parents by phone. During this initial contact, we conducted a brief screening to target children who met our language inclusion criteria, and confirmed that parents were willing to commit to three full day of recordings at home, including two weekdays and weekend.

Family information

All families consisted of one father (age range = 27-36 years, M = 36) and one mother (age range = 30-41 years, M = 35). Eight of the infants were first single-born children, ten had one older sibling, and two had two older siblings (sibling age range = 2-8 years; M = 4). These families were from mid- to high- socioeconomic backgrounds, with an average Hollingshead score of 52.2 (range: 31 - 66 out of a possible score of 66; Hollingshead, 1975). Most of the infants stayed at home with one or both of the primary caregivers (n = 16), while some of the infants stayed at home with a nanny during the weekdays (n = 2) or were enrolled in full-time daycare during the course of the study (n=3; note that many children in Montréal enrol in daycare around age 12 months, due to government parental leave policies).

Mothers

We interviewed the mothers of our participants to gain more information about their language background; note that self-report data were missing for one mother as she was not present during the interview. All of the mothers had knowledge of both French and English. A further five knew a third language, and three knew a fourth language, including Spanish (4), Portuguese (3), Arabic (1), German (1), Italian (1), and Kannada (1). Mothers had an average of 17.9 years of education (range = 11 – 23; SD = 3.1). Based on the Bilingual Dominance Scale (Dunn & Fox Tree, 2009), there was a range of English-dominant and French-dominant bilingual mothers in our sample, with a mean of -1.9 (SD = 11.3) and a range of -19 to +17 (possible scores = -30 to +30, with 0 indicating a completely balanced bilingual, -30 indicating an English-dominant bilingual, and +30 indicating a French-dominant bilingual. Our sample of mothers reported a wide range of rates of intra-sentential language mixing as measured by the Language Mixing Scale (Byers-Heinlein, 2012), with a mean of 12.4 (SD = 9.8) and a range of 0 to 30 (possible scores = 0 to 30, with 0 indicating no language mixing and 30 indicating frequent language mixing).

Fathers

We also interviewed the fathers of our participants; note that self-report data were missing for one father as he was not present during the interview. All of the fathers also had knowledge of both French and English; three fathers knew a third language, including Arabic (1), Portuguese (1), and Spanish (1). They had an average of 17.1 years of education (range = 14 – 22; SD = 2.3). Based on the Bilingual Dominance Scale (Dunn & Fox Tree, 2009), fathers in our sample were more French-dominant than English-dominant (*M* = 8.9, SD = 12.8, Range = -18 – 22). Fathers in our sample also reported a

wide range of scores in the Language Mixing Scale (Byers-Heinlein, 2012), with a mean of 9.3 (SD = 8.0) and a range of 0 to 23.

3.2.2. Procedure and measures

The study consisted of three sessions: the initial laboratory visit, the home recordings, and the final visit. During the family's initial laboratory visit, we conducted a comprehensive language environment interview with one *(n* = 17) or both (*n* = 4) caregivers (described in more detail below). Parents were then asked to do three full day recordings at home: two on the weekdays and one on a weekend day. All but two families were able to follow this schedule; one family did the recordings on one weekday and two weekend days, while another family did the recordings on three weekdays. At the end of each recording day, parents were asked to fill out a daily activity diary, detailing the infant's general activities throughout the day. Finally, we visited the family at their homes or at our laboratory to collect the recording devices and complete language and demographic questionnaires with one (n = 2) or both (n = 19) caregivers. We also conducted the i) Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld & Kaushanskaya, 2007), which assessed each parent's proficiency levels in each of their language, ii) the Bilingual Dominance Scale (BDS; Dunn & Fox Tree, 2009), which quantified their potential dominance in one of their languages, and iii) the Language Mixing Scale (LMS; Byers-Heinlein, 2012), which examined the extent to which each parent mixed their two languages when speaking to their child.

Reported measures of language exposure

Here, we examined two measures of reported language exposure: the *pre-study lifetime estimate*, and the *post-recording estimate*. For both cases, the estimates are

defined as the proportion of total words heard in language X out of total words heard in all languages.

The *pre-study lifetime estimate* was obtained during the interview portion of the initial laboratory visit. We asked parents a series of questions are typical of other language interview formats from other laboratories (Bosch & Sebastián-Gallés, 2001; DeAnda et al., 2016), following established practices for conducting the language interview (Byers-Heinlein et al., under review). To help parents recall what languages their child typically hears at home, we first asked them a range of language environment questions, including what languages each member of the household speaks to each other, what languages their child is exposed to in different contexts (television, play time, meal time, book reading, songs), and what languages other members of their family or community speak to them (grandparents, family friends, neighbours, day care teacher). We also asked them what family language practices they adopted at home, if any. Then, we asked parents to describe their child's typical day and estimate what languages their child hears on weekdays and weekends, on a month-to-month basis, from birth to the present day. We asked parents to consider only speech directed towards their infant, and not speech from the TV or radio nor overheard speech. During this process, we asked parents to consider any situations when the language proportion might be different, including whether one or both parents were at home or at work, whether they went on long trips abroad, or whether they had long-term visitors at their home. Finally, we asked parents for a global estimate of language exposure to French, English and any other present languages. We took the average of their month-to-month estimate and their global estimate to obtain the pre-study lifetime estimate of proportion exposure to each language.

The *post-recording estimate* was obtained during the final visit. We asked parents the following question: "Think about the days that you recorded. If you counted all the words your child heard in the recordings, what percentage do you think would be in each language?" Parents were given a chance to review the daily activity diary that they filled out at the end of each recording day. For each recorded day, parents gave an estimated percentage for English, French and any other languages.

Observed measures of language exposure

Parents were asked to complete three full days of recording using the LENA recording system. They were instructed to dress their child with a custom-made vest at the beginning of the day, and insert the recording device into the pocket of the shirt or vest. Parents were asked to carry on with daily activities and speak to their child as they would on a typical day. They were asked to keep the recording device running for the entirety of the day, until the recording device automatically stopped after 16 hours. To ensure that incidental study participants also provided consent, parents were required to ask other individuals around the infant to sign a consent form. They were also instructed to pause the recording if they were overhearing strangers' voices when outdoors, or to note down the time of day so that the relevant portion could be deleted for analyses. At the end of each recording day, parents were asked to complete a daily activity diary.

In total, the families produced 1008 hours of audio recordings (21 families X 3 days X 16 hours). The LENA algorithms estimate certain variables in the recording, including word counts from adults, frequency of infant vocalizations, amount of turn taking between the infant and an adult, amount of media sounds, and amount of noise in the background. To extract more information about the language input from the recordings, we constructed

a coding scheme that was inspired by the infant Social Environment Coding of Sound Inventory (SECSI; Ramírez-Esparza, García-Sierra, & Kuhl, 2014). First, we divided the recordings into 30-second segments via Audacity Software Version 2.1.1 (Audacity Team, 2014). We then matched these segments with the LENA-generated Adult Word Count using LENA software. We discarded all segments that had zero word counts from analyses. Then, trained research assistants listened to each segment and tagged them for social context (e.g., *Infant is alone*, *With one other individual*, or *With two or more individuals*), speaker context (i.e., who is speaking and to whom: *Mother, Father, Sibling, Infant, Other*), and language context (i.e., what language was being spoken: *English, French, Mixed, Unknown*). Based on pilot analyses, we determined that coding half of the recordings was sufficient to obtain reliable measures of the language proportion breakdown in the recording. Thus, we decided to code every other 30-second segment. Seven research assistants completed this time-intensive process. All research assistants were highly proficient simultaneous French-English bilinguals from Québec, Canada, and were undergraduate students majoring in Linguistics or Psychology; each completed a training file before coding the data files for this project. To calculate the language proportion in the recordings, we took the LENA-generated adult word counts that were tagged in each language and divided it by total adult word count tagged in both English and French. The resulting dataset is a rich description about the child's language experiences during the recordings.

3.3. Results

3.3.1. Parent-report versus observed measures

Language proportion

First, we examined the accuracy of parent reports in estimating the observed language exposure in the LENA recordings. Recall that we gathered two reported measures of language exposure: the *pre-study lifetime estimate*, and the *post-study estimate*. The observed measure here is the average language proportion across the three days of recording. Figure 3.1 plots the relation between reported and observed measures of language proportion.

The results show that parents' initial assessment of their child's language environment closely matched the language proportions heard in the recordings. Indeed, Pearson's correlation analysis revealed a positive and strong relation between the pre-study lifetime estimate and the observed measure [$r = .76$, $p < .001$]. Figure 3.1a shows that, apart from five participants, the dominant language (i.e., most-heard language, >50% exposure) according to parent report was also the dominant language heard in the recordings. The median absolute difference between these two variables was 14%. It is important to note that the pre-study lifetime estimate took into account the language proportion from birth to present day, while the observed measure only accounted for the present day's language proportion during the three recorded days. Given that the proportion of language exposure could change from birth to present day (e.g., if one caregiver goes back to work, if the child is in daycare, or if relatives visit for a long period of time), the calculation for the pre-study lifetime estimates may have included language proportions that were no longer current or relevant. Further, the pre-study lifetime estimates took into account the language proportion through a typical week (5 weekdays and 2 weekend days), while the recordings only involved two weekdays and one weekend day. Despite this difference in operationalization, the two variables were still closely

related, suggesting that parent-report measures of language exposure may be sufficiently stable for describing a child's current exposure at this age. Note also that the participating infants were relatively young and, in most cases, were at home with a primary caregiver. Thus, our group of infants may have experienced fewer changes either across different days of the week or across different months of their lives so far, as compared to older infants.

Analysis of the post-recording estimates also indicated that parents can reliably describe their child's language exposure from the recent past (see Figure 3.1b). Recall that, at the end of the study, parents were asked to estimate the language proportion of each day; we took the mean of these reported language proportions across the three days for each participant. Pearson's correlation analysis showed that this post-recording estimate was significantly correlated with the observed measure [$r = .78$, $p < .001$]. The median absolute difference between these two variables was 11%. In this analysis, both estimates are averaged across estimates of language proportion in two weekdays and one weekend day, reflecting a more valid comparison of reported versus observed measures of language exposure. One possible limitation is that parents may be hyperaware of the languages that they were speaking through the day, given that they knew this was the focus of the research. However, since parents were not told ahead of time that they would be asked this question, this decreases the likelihood that parents were consciously keeping track of the languages that they were speaking during the recordings.
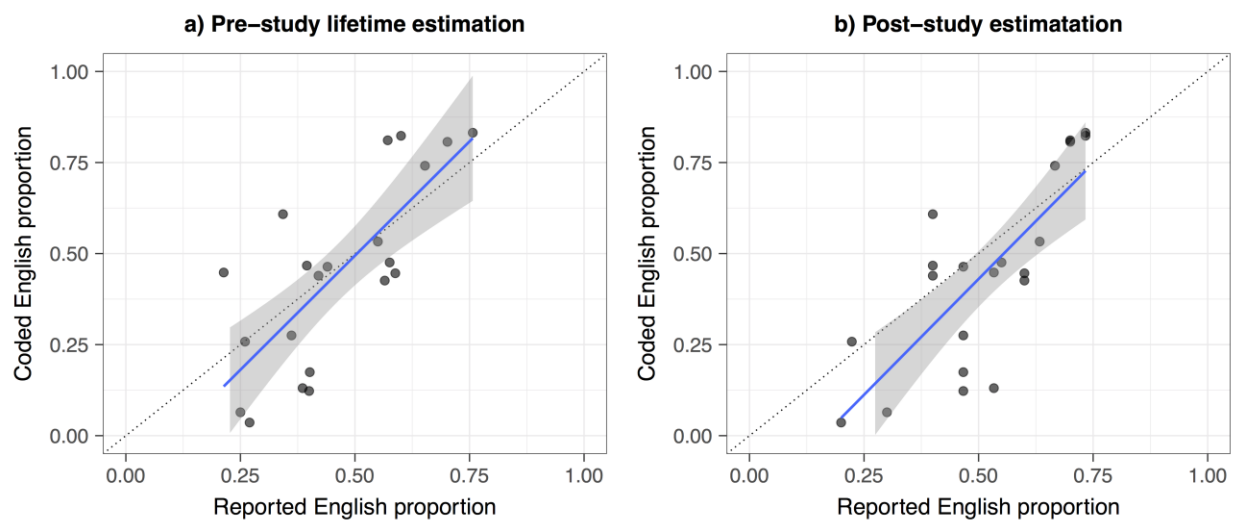
Figure 3.1. Relationship between reported and observed language exposure, where the reported estimate was the: a) Pre-study lifetime estimation, and b) Post-study estimate by day. Dotted line represents a perfect 1:1 match between variables on the x- and y-axis. The blue solid line represents the best-fitting regression line, and the shaded region represents the 95% confidence interval.

Language use

An important aspect of the language input to children is what languages they are hearing from their parents. Here, we evaluate the accuracy of parent reports of their own language use. During the pre-study language interview, we asked the present parent(s) to identify what languages both parents speak, as well as what languages they speak to their infant. Based on this information, we split the parents into those who reported speaking only one language (n = 16), and those who reported speaking both languages (n = 26) to their infant. Figure 3.2 plots group and individual data points of the language proportion that parents spoke to their child based on home recordings. All of the bilingual parents who reported speaking only one language to their infant were consistent (although not perfectly so) in using their target language (n = 15; Mean of using their target language = 97%; *SD* = 3%), while those who reported speaking both languages

used their dominant language more often (n = 27; Mean of using their dominant language = 84%; *SD* = 16%).

Note, firstly, that there is some overlap in language use between those who reported speaking only one language and those who reported speaking two languages, suggesting that this question by itself did not differentiate the two groups. Interestingly, all parents spoke both French and English to their children at least minimally, including those who reported speaking only one language to their child. This might reflect the highly bilingual nature of adults in Montréal.

Second, note that very few parents were balanced in their two languages when speaking to their 10-month-old: only 5 out of 42 parents spoke each language at least 25%-75% of the time. Interestingly, the five parents that tended to have the most balanced use of their two languages with their infant reported that they spoke more English than French, and they were more English-dominant (per their Bilingual Dominance Scale scores). This is consistent with the socio-cultural trend in nearby communities (Ottawa-Hull) that English-dominant bilinguals tend to switch their languages more often than French-dominant bilinguals in a French-dominant environment (Poplack, 1988).
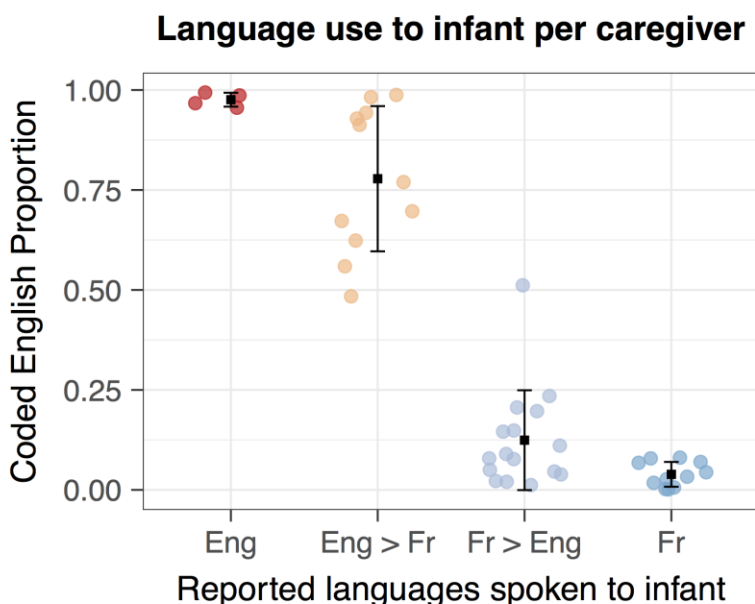
Language use to infant per caregiver

Figure 3.2. Reported and observed languages spoken to infant, where each dot represents one caregiver.

We also examined how well parents followed the family language policy that they implemented at home. Prior studies suggest that parents are not very good at strictly adhering to the one-parent, one-language strategy (De Houwer, 2016; Goodz, 1989), but it is not known how much parents deviate from an intended one-language use pattern. In our sample of families, four families explicitly reported using the one-parent, one-language strategy with their children (Table 3.1). For Participants 1 to 3, the father provided the English input to their child, while the mother provided the French input, while the reverse was the case for Participant 4. Consistent with findings from Goodz (1989), this table shows that parents who explicitly chose this language policy were relatively consistent in using their target language. All parents spoke their "target" language more than 93% of the time to their infant. Nevertheless, it is important to note that a one-parent, one-language policy does not necessarily translate to a balanced exposure to each

language. As shown on Table 3.1, a child's overall exposure to each language can also vary, depending on which parent spends more time with the child. Despite hearing one language from one parent each, one child (Participant 2) was exposed to English 67% of the time, while another (Participant 3) was exposed to English only 40% of the time.

Table 3.1. Observed languages spoken to the infant for families who report using a one-parent, one-language strategy at home

	English proportion, by caregiver		Overall exposure to English
	Mother	Father	
Participant 1	0.98	0.04	0.50
Participant 2	0.99	0.02	0.67
Participant 3	0.96	0.03	0.40
Participant 4	0.07	0.94	0.44

3.3.2. Nature of bilingual language exposure

Language consistency across different days

When assessing a bilingual child's language experiences, we typically estimate their overall language exposure to each language. But, how consistent is a child's exposure to each language across a child's day-to-day exposure to language? Even in parent-report measures, parents usually note differences between weekdays and weekends – often due to differences in child care (e.g., parents working on the weekdays so another caregiver is present). However, it is not well documented whether parents themselves are consistent in the languages that they speak to their infant from day to day.

We excluded two families (i.e., 4 caregivers, 2 infants) from these analyses, since they did not complete the requested two weekdays and one weekend day of recording.
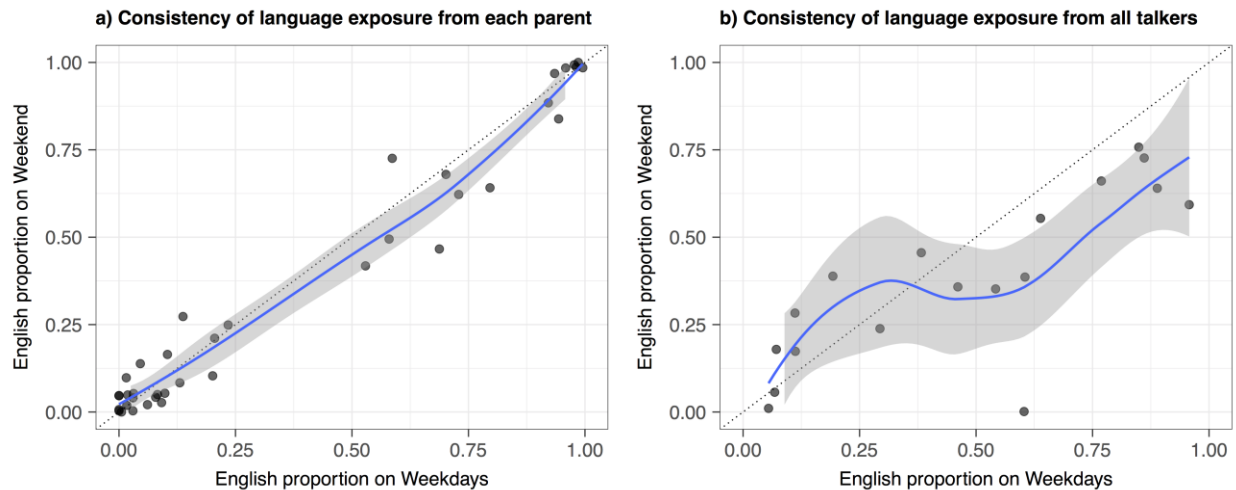

Figure 3.3. a) The relationship between infant's proportion exposure to English from each parent during the weekdays and the weekend, where each dot represents one parent; b) The relationship between infant's proportion exposure to English from all speakers during the weekdays and the weekend, where each dot represents one infant. Dotted line represents a perfect 1:1 match between variables on the x- and y-axis. The blue solid line represents the LOWESS line (Locally weighted scatterplot smoothing), and the shaded region represents the 95% confidence interval.

First, we examined the consistency of infant's language exposure from each parent across the different days. We computed proportions for each parent separately for weekends and weekdays (weekday scores averaged across the two recording days), plotted in Figure 3.3a. An examination of locally weighted regression curves (i.e., LOWESS: Locally Weighted Scatterplot Smoothing) suggested a monotonically increasing relationship between weekend and weekday input, which was not always linear. Thus, we examined this relationship using Spearman's rank correlation. Indeed, there was a robust consistency in the proportion that parents used each language to their

child on the weekdays versus weekend [r_s = .94, p < .001], showing that parents who spoke proportionally more English on weekdays also spoke proportionally more English on the weekends. The median absolute difference in language use proportion between weekdays and the weekend day is only 3.7% (range = .2 – 22.1%). Thus, when looking across weekdays and weekends, most parents were consistent in the languages they spoke to their infant.

Note, however, that even though parents are consistent in their language use when speaking to their infant, a child's overall input to each language depends on the amount of time spent with each parent speaking those languages. As previously mentioned, some of the infants in our sample were cared for at home by the primary caregiver, while the other caregiver was at work. Thus, we also examined infant's overall exposure to each language across the different days. Here, we considered only the speech directed towards the infant. Data analyses show a significant Spearman's rank correlation between an infant's language exposure on weekdays and the weekend day [r_s = .71, p < .001]. Nevertheless, note the wide individual variability shown in Figure 3.3b. The median absolute difference in language exposure between weekdays and the weekend day was 10.8% (range = 1.1 – 60.2%). These data remind us that bilingual infants' language experiences can vary widely by day especially when different caregivers are present, and they caution us about sampling bilingual infants' language experiences within a limited observation period.

Speech directed to infant versus to others

The auditory environment at home does not consist only of speech directed to infants, but also speech directed to other members of the family. While infant-directed

speech has been shown to be the central factor in how input supports children's language development (e.g., Ramírez-Esparza et al., 2014), there is some evidence that infants can also learn some aspects about language from overheard speech (Oshima-Takane, Goodz, & Derevensky, 1996; Shneidman, Buresh, Shimpi, Knight-Schwarz, & Woodward, 2009). In this analysis, we examined how parents differed in their language use when speaking to their infant versus other members of their family. Indeed, our sample of infants was exposed to more other-directed speech than infant-directed speech [V = 208, p < .001]. On average, they heard other-directed speech 1.66 times (*SD* = .82) more often than infant-directed speech. Only three infants heard more infant-directed speech than other-directed speech.
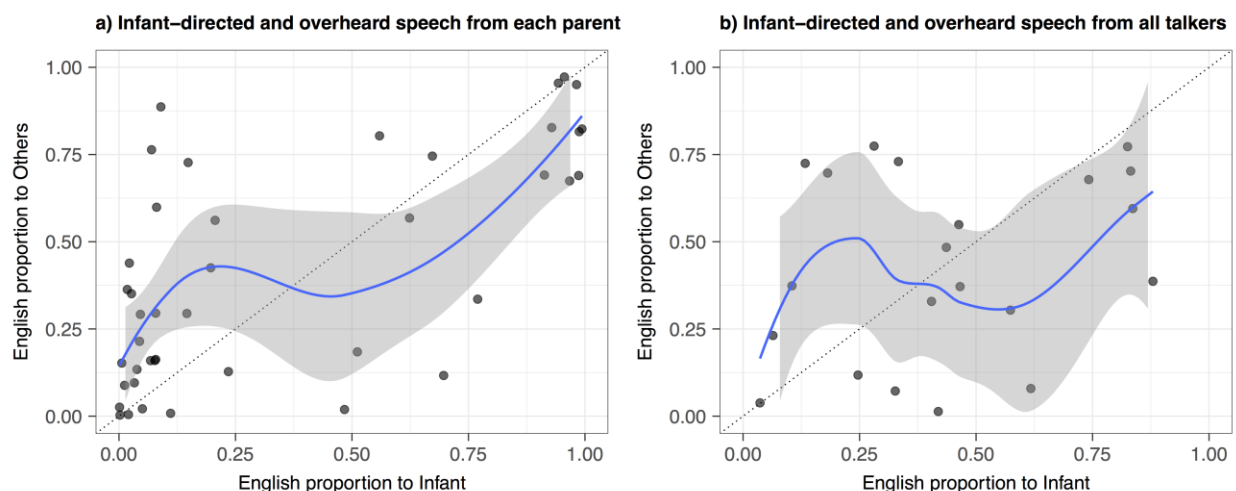

Figure 3.4. a) The relationship between infant's proportion exposure to English from each parent when speech is directed towards the infant versus directed towards other members of the household, where each dot represents one parent; b) The relationship between infant's proportion exposure to English from all talkers when speech is directed towards the infant versus directed towards other members of the household, where each dot represents one infant. Dotted line represents a perfect 1:1 match between variables on the x- and y-axis. The blue solid line represents the LOWESS line (Locally weighted scatterplot smoothing), and the shaded region represents the 95% confidence interval.

We examined whether parents changed the proportion of English vs. French used when speaking to their infant versus other members of the family. Figure 3.4a plots the relationship between proportion of English use by each parent when directed towards the infant or directed towards others. Spearman's rank correlational analyses suggests that parents are consistent in their language use regardless of the addressee [r_s = .65, p < .001]. Note, however, that this correlation is likely driven by parents who use only one language at home, both with their infant and with other adults (see Figure 3.4a). Indeed, there is large variability in whether or not parents changed their language proportion. The median absolute difference in language use by addressee is 17.0% (range = 1.0% - 79.7%). These data highlight the possibility that even though infants are only hearing their parent speak one language to them, they may be hearing the same parent speak other languages to other members of the household.

Indeed, when considering the input landscape more broadly, our sample of infants tended to hear their language in different proportions when we compare all infant-directed and other-directed speech in their input. Figure 3.4b plots the differences in language exposure between these two types of speech per infant. Spearman's rank correlational analyses reveal no significant relationship between language exposure via infant-directed speech and other-directed speech [r_s = .28, p = .22]. The median absolute difference in language use by addressee is 24.1% (range = 1.7 – 59.1%). This motivates deeper consideration how these differences might affect early speech perception, especially since young infants hear more overheard speech than infant-directed speech.

3.4. Discussion

The general aim of the current study was to explore the quantitative nature of dual language input to bilingual infants. To do so, we collected day-long recordings from French-English bilingual families in Montréal, along with parent-report measures of family language use. This study has two main contributions. First, this study provides support for using parent reports to assess a bilingual child's language environment. Second, this study shows – via a data-centered approach – the wide heterogeneity in bilingual infants' experiences even within the same community. These findings provide new insights on how infants in bilingual environments experience language input, which can inform both research and clinical assessments with bilingual infants.

Assessment of infants' language environments

First, our study reveals that language interviews can elicit reliable quantitative information about a bilingual infant's language environment. Specifically, parent-report measures of their child's exposure to each language correspond closely to language proportions observed in naturalistic daylong recordings at home. This finding expands on previous research (Marchman et al., 2016) by showing this pattern in both a predictive and retrospective direction. Our approach was to conduct a language interview both before and after the recordings were made (i.e., pre-study lifetime estimates and post-study estimates, respectively), and we showed that both types of parent-report measures were tightly coupled with infants' actual language experiences. By showing this relationship in both directions, we can deduce that i) the recordings conducted at home are representative of a child's typical day, and that ii) parents can accurately recall the language proportion of days that have passed.

We nonetheless still observed some small discrepancies between the parent-report and observed measures of language proportion. Some of these discrepancies are likely due to non-systematic measurement error, which is an inherent part of the measurement process. However, it is also possible that parent reports are affected by reporting-related factors. For example, caregivers' judgment of their child's language environment may be clouded by their own experiences with language. Indeed, our data shows a stronger numerical correlation between the post-study estimate and the observed measures when considering all speech heard by the infant (i.e., including speech by adults to other adults or children in the household) [r(19) = .87, p < .001], versus when considering just speech that was directed to the infant [r(19) = .78, p < .001]. Future confirmatory research is needed to corroborate these differences, and to examine what other external biases are implicated in individuals' ability to track their language use.

Our findings also show that parents are quite reliable at describing their own language use to their infants. While all bilingual parents used both languages when speaking to their infant even minimally, those who reported speaking one language used their dominant language proportionally more than those who reported speaking both languages. Further, the families who explicitly reported using the "one-parent, one-language" strategy generally showed this pattern during the recordings, although in our sample all parents used both languages at least to a minimal degree.

While much of the data in this paper concerns the *proportional* amount of language exposure, it is important to also consider *absolute* amount of language exposure. Prior studies have shown that these two variables do not always align with one another (Marchman et al., 2016). In our own dataset, the proportional measures of language

exposure did correlate with the absolute measures of input for both French [r(19) = .62, p = .003] and English [r(19) = .62, p = .003]. However, while relatively strong, these correlations are far from perfect. This means that two children with similar relative amounts of exposure to each language might have quite different absolute amounts of exposure. As an extreme example, one child in our sample who was reported to hear English 57% of the time heard 10,098 English words per day on average, while another child who was reported to hear English 59% of the time heard only one-tenth the absolute input: 1,184 English words per day on average. Indeed, some studies have shown that the absolute amount of language input is a better predictor of language outcomes than proportional measures of language exposure (De Houwer, 2011; Marchman et al., 2016). Nonetheless, competition models of bilingualism point to the importance of relative exposure as well (Hernandez, Li, & MacWhinney, 2005). Future studies should examine how these two types of input measures interact with proficiency levels in both languages.

Variability in bilingual exposure

Many studies that examine infants' language experiences focus on the quantity of speech directed to infants. In research with bilingual infants, researchers most often measure infants' proportional exposure to each language. However, the ways in which infants accrue this experience vary considerably. In the current paper, we provide data showing variability in bilingual exposure in a small group of bilingual infants from the same community. While this is not a novel insight on the bilingual experience, our study describes this variability in naturalistic home recordings in a more direct and detailed way, and reminds us to consider these individual differences when examining bilingual infants in group-level comparisons with monolingual infants.

Indeed, even with our limited sample size, we observed wide array of bilingual caregiving behaviors and approaches. With regards to speech directed towards their infant, some families reported that both parents spoke both languages (n = 9), others reported that one parent spoke both languages while the other spoke only one of the two languages (n = 8), and some reported using the one-parent, one-language strategy for their child (n=4). These choices may be due to strong individual differences in parents' own bilingual histories. For example, self-report data indicated that our sample of parents differed in their proficiency levels for each language, in their comfort and preference for using each language, and in their amount of daily exposure to each language. These language characteristics are non-trivial, as previous research has shown that they can have an impact on the quality of their speech, and as a result, on their infant's speech processing abilities (e.g., van Heugten & Johnson, 2017). Interestingly, although most parents were fluently bilingual, very few parents were actually using both languages in balanced proportions when speaking to their 10-month-old. That is, most parents tended to speak their dominant language most of the time. It would be of interest to examine how these language proportions might change over time, as the child begins to understand and produces words in both of their languages. It is possible that parents are more likely to stick to one language when interacting with younger infants, but that this may change over time to match children's proficiency in comprehending and producing words in each language.

We also examined how infants experienced their parents' language use in their natural environment. Our data showed that, across the three recording days, parents were fairly consistent in what languages they used when speaking to their child. However,

many parents were not consistent in the languages that they spoke to different members of the household (i.e., infant versus others). For example, a caregiver might consistently use Language A with their infant, but use both Language A and B while speaking to their partner. Critically, we found that infants' proportional exposure to each language differed when considering infant-directed versus other-directed speech.

These findings motivate future research to examine how overheard speech might affect infants' speech perception skills. To date, the focus has largely on infant-directed speech, especially given the large body of work showing its importance on different aspects of speech perception and language outcomes (see Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015). However, it is also an important question of how overheard speech might also contribute to infants' language acquisition (e.g., Sperry, Sperry, & Miller, 2018). Our data reveals that infants "hear" more speech from their parents addressed to other members of their household than to themselves (with the caveat that the recordings cannot explicitly tell us whether infants are paying attention to the overheard speech). Certainly, some studies indicate that young monolingual children can learn aspects of their native language from overheard speech (e.g., Shneidman et al., 2009). Thus, it would be interesting to consider how overheard speech (in addition to infant-directed speech) might play a role in bilingual infants' speech perception and processing abilities.

Limitations and Conclusions

There are some limitations to the generalizability of our findings. First, we acknowledge that even though we sampled three full days of recordings (including two weekdays and one weekend day), these may still not be representative of infants' full range of language experiences. Some of the parents did note that they felt more inclined

to stay at home during the recordings so that they did not need to worry about obtaining consent from other individuals who might happen to get recorded. This is an inherent challenge in collecting daylong recordings from families, and motivates more in-depth examination of the bilingual input in different contexts.

Second, our sample of French-English bilingual families from Montréal may not be wholly representative of all bilingual families' ability to report their child's language environment. Indeed, the topic of *language* is prominent in the social and political context of Montréal; thus, it may be that caregivers in Montréal may be more aware of the languages being spoken to and around their child. Further, the current experiment required parents to be willing to record two weekdays and one weekend of their daily lives. As such, our sample of families were of mid- to high-SES, and most of the infants were being cared for at home (i.e., not at daycare). While this is a typical set-up for families in Montréal with children of this age (given parental leave policies in Canada), the study may still have selected for families who would have more opportunities to be aware of their child's language environments. Indeed, parents who care for their child at home spend more time with their children, and would be more aware of what languages their child hears from the different speakers in their lives. Thus, these results might not generalize as well to those who are enrolled in daycares, or to those who spend more time with other individuals (including grandparents and older children). Nevertheless, even within our own sample, the primary caregiver (who tended to be the person to participate in the language interview) was not always around the child throughout the day, suggesting that caregivers can also estimate what their child hears from other speakers.

Similarly, in interpreting the current data on variability of bilingual experiences, we caution in generalizing to other bilingual communities outside of Montréal. Indeed, bilingual communities around the world differ in many socio-cultural dimensions, and various sociolinguistic differences have been shown to play a role in language acquisition (Smithson, Paradis, & Nicoladis, 2014; Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007). Montréal is socio-linguistically unique, as both French and English are of high status, and most adults are bilingual. Indeed, some of our findings may be directly tied to the sociolinguistic context in Montréal: for example, since both French and English are widely spoken between adults in Montréal, infants may be more likely to overhear two languages in adult conversations, compared to infants from other environments where one language is more dominant. As our field collects more daylong recordings of infants' language experiences from other bilingual communities, it would be interesting to compare and contrast infants' bilingual experiences. Identifying different "bilingual profiles" would be a critical step towards understanding how different language experiences affect language acquisition.

In sum, the present study examined the quantitative nature of dual language exposure in young bilingual infants. Specifically, it provides support for conducting language interviews for assessing a child's language environment, but it also motivates the need to examine the bilingual experience beyond proportional measures. Indeed, our findings highlight the individual differences in bilingual experience, even in a small sample of bilingual infants learning the same two languages in the same city. Examining these different types of bilingual input is important for our understanding of input effects in bilingual acquisition.

Acknowledgments

3.5. References

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*. https://doi.org/10.1080/15248372.2010.516420

Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2016). Early Prosodic Acquisition in Bilingual Infants: The Case of the Perceptual Trochaic Bias. *Frontiers of Psychology*, 7, 210. https://doi.org/10.3389/fpsyg.2016.00210

Bijeljac-Babic, R., Serres, J., Höhle, B., & Nazzi, T. (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLoS One*, *7*(2), e30843. https://doi.org/10.1371/journal.pone.0030843

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2–3), 217–243. https://doi.org/10.1177/00238309030460020801

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

Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, *35*(2), 225–241. https://doi.org/10.1017/S0142716412000379

Byers-Heinlein, K. (2012). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism*, *16*(1), 32–48. https://doi.org/10.1017/s1366728912000120

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

Byers-Heinlein, K., Schott, E., Gonzales-Barrero, A. M., Brouillard, M., Dubé, D., Jardak, A., … Tamayo, M. P. (under review). Best practices for using parent report to measure bilingual infants' language environments.

Carroll, S. E. (2015). Exposure and input in bilingual development. *Bilingualism*, 1–14. https://doi.org/10.1017/s1366728915000863

Core, C., Hoff, E., Rumiche, R., & Señor, M. (2013). Total and Conceptual Vocabulary in Spanish–English Bilinguals From 22 to 30 Months: Implications for Assessment. *Journal of Speech Language and Hearing Research*, *56*(5), 1637. https://doi.org/10.1044/1092-4388(2013/11-0044)

De Houwer, A. (2007). Parental language input patterns and children's bilingual use.

Applied   Psycholinguistics,   28(3),   411–424.
https://doi.org/10.1017.S0142716407070221

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

De Houwer, A. (2016). Bilingual language input environments, intake, maturity and practice. *Bilingualism*, 1–2. https://doi.org/10.1017/s1366728916000298

De Houwer, A., & Bornstein, M. H. (2016). Bilingual mothers' language choice in child-directed speech: continuity and change. *Journal of Multilingual and Multicultural Development*, 1–14. https://doi.org/10.1080/01434632.2015.1127929

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Frienda, M. (2016). The Language Exposure Assessment Tool: Quantifying Language Exposure in Infants and Children. *American Journal of Speech-Language Pathology*, *25*, 1–15. https://doi.org/10.1044/2016

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

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

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

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

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H Brookes Publishing.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2005.03.003

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

Hollingshead, A. (1975). Four factor index of social status. Unpublished manuscript, Yale University, New Haven.

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism*, *17*(1), 189–202. https://doi.org/10.1017/s136672891300014x

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3), 159–207. https://doi.org/10.1006/cogp.1999.0716

King, K. A., Fogle, L., & Logan-Terry, A. (2008). Family language policy. *Linguistics and Language Compass*. https://doi.org/10.1111/j.1749-818X.2008.00076.x

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

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2016). Caregiver talk to young Spanish-English bilinguals: comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*. https://doi.org/10.1111/desc.12425

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

Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: Do secondborn children learn from overheard speech? *Child Development*, *67*(2), 621–634. https://doi.org/DOI 10.1111/j.1467-8624.1996.tb01755.x

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41–58. https://doi.org/10.1017/S0142716400009863

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

Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2016). Segmenting words from fluent

speech during infancy - challenges and opportunities in a bilingual context.

Developmental Science. https://doi.org/10.1111/desc.12419

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative

vowel contrasts. *Journal of Experimental Psychology: Human Perception and*

Performance, 20(2), 421. https://doi.org/10.2307/1129499

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective

attention to the mouth of a talking face. *Psychological Science*, *26*(4), 490–498.

https://doi.org/10.1177/0956797614568320

Poplack, S. (1988). Contrasting patterns of code-switching in two communities. *Aspects*

of Multilingualism, 51–77.

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

Sebastián-Gallés, N., & Bosch, L. (2002). Building phonotactic knowledge in bilinguals:

Role of early exposure. *Journal of Experimental Psychology: Human Perception and*

Performance, *28*(4), 974–989. https://doi.org/10.1037/0096-1523.28.4.974

Shneidman, L. A., Buresh, J. S., Shimpi, P. M., Knight-Schwarz, J., & Woodward, A. L.

(2009). Social Experience, Social Attention and Word Learning in an Overhearing

Paradigm. *Language Learning and Development*, *5*(4), 266–281.

https://doi.org/10.1080/15475440903001115

Smithson, L., Paradis, J., & Nicoladis, E. (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Bilingualism*, *17*(4), 810–821. https://doi.org/10.1017/S1366728913000813

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2018). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*. https://doi.org/10.1111/cdev.13072

Statistics Canada (2016). Focus on Geography Series, 2016 Census. Statistics Canada Catalogue no. 98-404-X2016001. Ottawa, Ontario.

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition*, *108*(1), 232–242. https://doi.org/10.1016/j.cognition.2007.12.013

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*. https://doi.org/10.1111/desc.12456

Theunissen, N., Vogels, T., Koopman, H., Verrips, G., Zwinderman, K., Verloove-Vanhorick, S., & Wit, J. (1998). The proxy-problem: child report versus parent report in health-related quality of life research. *Quality of Life Research*, *7*(5), 387–397. https://doi.org/10.1023/A:1008801802877

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

Titone, D., & Baum, S. (2014). The future of bilingualism research: Insufferably optimistic

and replete with new questions. *Applied Psycholinguistics*, *35*(5), 933–942. https://doi.org/10.1017/s0142716414000289

Vagh, S. B., Pan, B. A., & Mancilla-Martinez, J. (2009). Measuring growth in bilingual and monolingual children's english productive vocabulary development: The utility of combining parent and teacher report. *Child Development*, *80*(5), 1545–1563. https://doi.org/10.1111/j.1467-8624.2009.01350.x

van Heugten, M., & Johnson, E. K. (2017). Input matters: Multi-accent language exposure affects word form recognition in infancy. *The Journal of the Acoustical Society of America*, *142*(2), EL196-EL200. https://doi.org/10.1121/1.4997604

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

Weisleder, A., & Fernald, A. (2013). Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: first steps in perception and comprehension. *Trends in Cognitive Sciences*, *12*(4), 144–151. https://doi.org/10.1016/j.tics.2008.01.008

Preface to Chapter 4

The study in Chapter 3 revealed that caregivers are sufficiently reliable at estimating their bilingual infants' exposure to each of their languages, providing support for using caregiver reports as a measure of infants' language experiences. Further analyses show that bilingual infants experience bilingualism in various ways. The next step of the dissertation was to examine how this bilingual experience affects their language skills.

In Chapter 4, we examined how language experiences affect an important aspect of speech processing: word segmentation (i.e., the ability to recognize word forms in connected speech). To assess infants' language experiences, we used the language interviews described in Chapter 3. Further, we administered the Language Mixing Scale (Byers-Heinlein, 2011) for bilingual caregivers. To examine infants' word segmentation skills, we used a novel variant of the Headturn Preference Procedure (HPP) that examines segmentation in a dual-language stream. In Experiment 1, we tested monolingual infants' segmentation abilities in their native and non-native language; and in Experiment 2, we tested bilingual infants' segmentation abilities in both of their native languages. These data allowed us to examine the relationship between infants' language experiences and speech processing skills.

CHAPTER 4. Segmenting words from interleaved bilingual speech during infancy

4.1. Introduction

At what point can infants recognize words in sentential contexts? Prior research shows that infants begin to do so in their native language between 6- to 12-months of age, and that this developmental timeline is consistent for monolingual learners of various languages and dialects, including *English* (Floccia et al., 2016; Jusczyk, Houston, & Newso me, 1999), *French* (Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006; Polka & Sundara, 2012), *Dutch* (Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000), *German* (Hohle & Weissenborn, 2003), *Spanish* and *Catalan* (Laura Bosch, Figueras, Teixidó, & Ramon-Casas, 2013). But, what if the speech stream includes more than one language? How do infants deal with the challenge of segmenting words in two languages within the same discourse? Indeed, bilingualism is surging in many parts of the world, including in Canada (Statistics Canada, 2016). More and more infants are being exposed to a second language – either at home or at daycare, through the media, or through daily exposure in multilingual cities. In the current paper, we explore how language experience plays a role in young infants' ability to segment words in bilingual speech. Understanding how infants from various language backgrounds process speech in bilingual environments is critical to our understanding of infants' processing capacities and language development trajectories.

Consider, first of all, how infants face the challenging task of processing speech. Infants can recognize highly frequent and meaningful words when they are spoken in *isolation*, for example, their own names by 4 months of age (Mandel, Jusczyk, & Pisoni,

1995) and socio-pragmatic words used in daily routines by 9 months of age (e.g., "night night" and "peekaboo"; Syrnyk & Meints, 2016). However, infants do not typically hear words in isolation, but rather in multiword utterances (Brent & Siskind, 2001; Van de Weijer, 1999). Indeed, speech is naturally continuous, and word boundaries in speech are not reliably tagged by overt acoustic markers such as pauses. Unlike adults, young infants cannot rely much on lexical knowledge to identify the beginning and end of a word. Thus, the young language learner must make use of bottom-up cues to identify words form candidates in the speech stream. Indeed, researchers have discovered that certain regularities exist in the structure of a language (e.g., its prosodic patterns, phonoatactic regularities and the distributional properties of words; Jusczyk, 1999) that can act as cues for the beginning or end of a given word. Laboratory studies with both artificial and natural languages confirm that young infants can indeed exploit these regularities to find word forms in a speech stream (e.g., Jusczyk et al., 1999; Saffran, Aslin, & Newport, 1996), with some cues more strongly weighted for use during segmentation (Johnson & Seidl, 2009; Sandoval & Gómez, 2016). Cross-linguistic studies indicate that many of these cues are language-specific: that is, some cues are expressed differently depending on the infant's ambient language (Cutler, Mehler, Norris, & Segui, 1986).

Take, for example, how infants learn and use prosodic patterns to segment words in English and French – the languages of interest in this study. English is a stress-timed language that involves alternating strong syllables (i.e., longer duration, higher pitch and amplitude) and weak syllables (i.e., shorter duration, lower pitch and amplitude). The majority of bi-syllabic content words in English have a trochaic (STRONG-weak) stress pattern, for example as in the word CAN-dle (Cutler & Carter, 1987), and infants are highly

sensitive to this stress pattern of English words from early on. Newborn infants can discriminate between trochaic (e.g. DOC-tor) and iambic (weak-STRONG; e.g., gui-TAR) word stress patterns (Sansavini, Bertoncini, & Giovanelli, 1997), and between 6 and 9 months of age, English-monolingual infants develop a listening preference for the more frequently-occurring trochaic words over iambic words (Jusczyk, Cutler, & Redanz, 1993). Given these sensitivities, researchers hypothesized that English-monolingual infants learn to track stressed syllables and extract trochaic word forms in connected speech.

Jusczyk and colleagues (1999) provided strong support for this hypothesis with English-monolingual infants by implementing what is now the widely used procedure for testing infant word segmentation in natural speech. In this two-phase task, implemented via the Headturn Preference Procedure (HPP), infants are familiarized with target words in isolation (Familiarization phase) and then presented passages containing each target word and passages containing a novel word (Test phase). The test phase is an opportunity to assess infant listening preference; longer listening times to the familiar word passages versus the novel word passages is taken as evidence that the infant detected the word form in connected speech. In another common variant of the task, infants hear target passages during the familiarization phase and isolated target and novel words during the test phase. In a series of fifteen experiments using these two variants, Jusczyk and colleagues (1999) showed that English-learning 8-month-olds are indeed sensitive to word stress patterns and appear to extract trochaic patterns from fluent speech. At this age, infants even mis-segmented trochaic patterns that are not real words; for example, infants segmented "tar is" as a unit following familiarization with a speech stream that contained the words "...the gui**tar is**…". Overall, their findings

provided evidence that 8-month-old, English-monolingual infants segment word forms that match the prosodic pattern typical of their native language.

A language-specific pattern is also observed with infants learning French (Nazzi et al., 2006). Unlike English, French is a syllable-timed language that does not use stress to distinguish between different words. Nonetheless, stress does mark phrase boundaries in French; that is, the final syllable in a phrase tends to be stressed (e.g., *Je joue de la guiTARE; Dahan & Bernard, 1996)*. Research with French adults reflects these rhythmic properties: they are less attentive to word stress patterns compared to English adults (Dupoux, Peperkamp, & Sebastián-Gallés, 2001) and track syllable co-occurrence patterns, rather than stress patterns, when segmenting connected speech (Cutler et al., 1986). Although French-monolingual infants do not show a preference for either iambic or trochaic words, they are able to discriminate between the iambic and trochaic stress patterns by 6-months of age (Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009). Like their English-monolingual peers, French-monolingual infants show language-specific patterns of word segmentation abilities. Findings reported by Nazzi and colleagues (2006) reveal syllable-tracking segmentation in Parisian-French learning infants. Polka and Sundara (2012) also observed language-specific patterns when testing segmentation of bi-syllabic words in monolingual infants learning either English or French in Canada. Specifically, at 8 months of age, French-monolinguals segmented iambic words in French, but not trochaic words in English, while English monolingual showed the reverse pattern (Polka, Orena, Sundara, & Worrall, 2017; Polka & Sundara, 2012). This outcome is likely due, at least in part, to differences in the basic prosodic patterns of English and French, and is not completely driven by familiarity with the native language phonology.

Consistent with this view, there is evidence that infants can segment words in different dialects or languages so long as the target word forms follow the rhythmic pattern prevalent in their native language. For example, Canadian French infants can segment iambic words in European French (Polka & Sundara, 2012). Likewise, English infants can segment trochaic words in other stress-timed languages (*Dutch:* Houston et al., 2000), and even in syllable-timed languages (*Italian:* Pelucchi, Hay, & Saffran, 2009). Thus, while stress may not be an important feature distinguishing word forms in French, it appears that prosody guides the segmentation of words in French infants.

The example of English and French above illustrates how the cues for segmentation may manifest in different ways for different languages, and how monolingual infants quickly learn how word forms are established in their ambient languages. How might this process be different for bilingual infants? They face at least four challenges. First, bilingual infants have to learn how word forms vary in two different languages instead of one, and this may be more challenging if, as discussed above, the cues for segmentation are different for each of the languages that they are learning. Indeed, prior work in the phonological domain suggests that bilingual infants may require more time to tune into their native language phonetic categories compared to monolingual infants, especially when there are conflicting phonetic patterns across their two native languages (Bosch & Sebastian-Galles, 2003; but, see Albareda-Castellot, Pons, & Sebastián-Gallés, 2011). Second, bilingual infants tend to receive less input in each of their native languages compared to what a monolingual receives in their single language (e.g., Garcia-Sierra, Ramírez-Esparza, & Kuhl, 2016), which in turn gives them less data in each language to discover the cues. In other words, bilingual infants may have less

opportunities to learn about the language-specific cues available for segmentation in each language, compared to monolingual infants. Third, it is likely that bilingual infants will receive some of their input as accented speech (e.g., hearing French words produced with English speech pattern), which may affect their ability to learn the correct cues in each language. Finally, the two languages are not always neatly separated in a child's input (Byers-Heinlein, 2012). Although in some cases the two languages are being used by different caregivers (e.g., dad speaks one language, while mom speaks the other language), in other cases the same caregiver is speaking both languages (e.g., mom speaks both languages). Both types of language mixing make the input noisier, and this larger amount of variability in speech may make it difficult for infants to track and attach specific cues to their respective languages.

A small number of studies have recently examined infants' word segmentation abilities in two languages. One study by Bosch and colleagues (2013) examined word segmentation in bilingual infants acquiring Spanish and Catalan – both syllable-timed languages. Both 6- and 8-month-olds showed positive evidence of word segmentation: Spanish-dominant bilingual infants were able to segment a target Spanish monosyllabic word from Spanish passages, and the same pattern was found with Catalan-dominant bilingual infants tested on Catalan materials. Like their monolingual peers, bilingual 6-month-olds showed the expected familiarity preference for the target words typically found in segmentation tasks, whereas 8-month-olds showed a novelty preference. This shift in preference direction may point to an increased efficiency in segmenting monosyllabic words with age (consistent with the model proposed by Hunter & Ames, 1988; but see Bergmann & Cristia, 2015). Overall, these findings suggest that bilingual infants follow

the same developmental trajectory as monolingual infants. However, note that Bosch and colleagues only tested their bilingual infants in their dominant language; thus, it is not clear from this study whether bilingual infants are able to segment words in both of their native languages. Another study by Singh and Foong (2012) was able to tackle this issue with English-Mandarin infants. They tested infants as young as 7.5 months of age in both of their native languages in separate test sessions: most infants did the two language sessions in direct succession (n=39), while some completed the second language session within five days of the first session (n=9). Infants were familiarized with monosyllabic target words, then tested on their recognition of those words in a sentential context. Results indicated that English-Mandarin infants were able to segment in both of their native languages by 7.5 months of age. Taken together, these studies suggest that bilingual infants follow the same developmental trajectory as monolingual infants in segmenting *monosyllabic* words from the speech stream.

A recent study by Polka and colleagues (2017) examined how bilingual infants might segment *bi-syllabic* words from the speech stream – a more challenging task that requires more linguistic knowledge. Importantly, segmenting monosyllabic words requires no knowledge of word stress patterns and thus, provides a limited view of word segmentation processes. For this reason, much of the initial work showing language-specific effects was focused on bi-syllabic segmentation (e.g., Jusczyk et al., 1999). Polka and colleagues (2017) tested 8-month-old infants on a dual-language segmentation task where English and French were tested in succession within a single test session. Under this protocol, results confirmed that 8-month-old, English and French monolingual infants could segment words in their native language, but not in a non-native language,

regardless of which language was tested first. However, French-English bilingual infants only showed successful segmentation for French, and only when it was tested first. Nevertheless, segmentation success was observed in English when separate groups of bilingual infants were tested in English and in French within the standard HPP task (i.e., familiarized with two target words, then tested with passages). Importantly, compared to the dual-language task, the standard task provided more speech exposure and no language switching occurred. Together, these findings suggest that bilingual infants are able to segment words in both of their languages but that performance is task-dependent. To segment in both languages, bilingual infants may need more exposure to sentences before being able to segment target words from those sentences, as well as unambiguous information on the nature of the input stream (i.e., whether the talker is speaking English, French, or both). These findings imply that bilingual infants' segmentation processes are less stable or involve more perceptual adaptation compared to their monolingual peers. In sum, there is still limited direct evidence showing that bilingual infants can segment bi-syllabic words in both of their native languages.

The present study aims to expand on this literature by providing more data about the speech segmentation capacities of infants in a bilingual context. Specifically, we investigated how language experience affects word segmentation when infants process two languages that are consistently inter-mixed. Prior work has only tested dual-language segmentation in successive language-blocked phases (Polka et al., 2017; Singh & Foong, 2012). However, languages are not always neatly separated by blocks in natural bilingual environments. Indeed, bilingual speakers often engage in frequent switching between their two languages in ongoing discourse, especially when both languages are present in

the environment (Heredia & Altarriba, 2001). This phenomenon is not limited to interactions between adults. In some cases, bilingual parents engage in language mixing with their children (Byers-Heinlein, 2012). Even infants growing up in monolingual homes can sometimes be exposed to bilingual speech (e.g., infants living in multilingual or multicultural cities). Thus, it is important to consider how infants might segment words when the incoming speech is language-mixed. By examining how monolingual and bilingual infants segment words in bilingual speech, we can gain more insight into infants' developmental trajectory and processing capacity for word segmentation.

To do so, we devised a new dual-language segmentation task in which the infant encounters two languages that are mixed throughout the test session – a context which likely occurs in a real interaction with a bilingual caregiver. This new mixed dual-language segmentation task differs from earlier language-blocked tasks in two ways. First, in the new mixed dual-language task, the infant encounters speech that mixes French and English (spoken by the same talker) throughout the task. Second, the initial familiarization phase is also longer in the new task (2 minutes of mixed language passage) compared to the blocked task (45 seconds per language). To accommodate the longer familiarization phase, we implemented a baby-friendly familiarization phase that exposes the infant to the passages in a passive manner (as in Bernard & Onishi, submitted). In other words, the auditory stimulus was played in the background, and this process was not infant-controlled. These changes provide the infant ample opportunity to process the passages that clearly specify a bilingual input. We tested French- and English-monolingual infants in Experiment 1, and French-English bilingual infants in Experiment 2.

The aim of testing monolingual infants in Experiment 1 was to validate the new mixed dual-language segmentation task. Specifically, we wanted to ensure that the findings conform to prior predictions about the language-specific nature of infant segmentation performance. Further, it is an important consideration, when testing bilinguals in two languages, to ensure that the task is equally sensitive in both languages and that any differences in performance for the two languages can be attributable to participants' processing skills and not to linguistic differences in stimuli or task difficulty. Since we planned to test both 8- and 10-month-old bilinguals in Experiment 2, we also examined both age ranges for monolingual infants in Experiment 1. For practical reasons[2], we recruited 8-month-old French- and English-monolingual infants and 10-month-old French-monolingual infants. With these groups, we were able to conduct both language and age comparisons (i.e., *Language comparison:* 8-month-old French vs. 8-month-old English; *Age comparison:* 8-month-old French vs. 10-month-old French).

If monolingual infants are unaffected by the mixed language context, then they should show language-specific patterns of segmenting in their native language, but not in the other unfamiliar language. However, it is possible that these infants may show different segmentation patterns in a bilingual discourse, especially given that they are not regularly exposed to this language context. On one hand, the extended exposure to mixed language passages may help bootstrap monolingual infants' ability to segment bi-syllabic words in the non-native language. That is, it is possible that hearing repeated words in the native stream might alert infants to also pay attention for repeated words in the non-

---

[2] It is more difficult to recruit English-monolingual infants in Montréal; thus, we decided to do the language comparison at one age (i.e., 8-month-olds), and the age comparison for one language (i.e., French).

native stream. On the other hand, the mixed language passages might perplex monolingual infants, as they may not be accustomed to hearing two languages from the same person. Thus, the language switches may distract them from segmenting words from either language stream.

4.2. Experiment 1: Methods

4.2.1. Participants

We recruited three groups of infants from Montréal, Canada: 8-month-old English-learning infants (n=16; Mean age = 246.5 days; Age range = 231 – 267 days; 10 boys), 8-month-old French-learning infants (n=16; Mean age = 244.9 days; Age range = 228 – 257 days; 8 boys), and 10-month-old French-learning infants (n=16; Mean age = 309.7 days; Age range = 297 – 324 days; 8 boys). Parent-reported measures indicated that infants received at least 90% exposure to their respective home language. An additional twenty infants were tested but excluded from the final sample (four 8-month-old English infants, six 8-month-old French infants, and ten 10-month-old French infants) because of fussiness (13), language background criteria exclusion (2), and technical error (5). Further, three infants were replaced since they were identified as outliers using a specific pre-established criterion: specifically, we excluded infants when their segmentation index (listening time during test trials minus control trials) for one of the languages was more than two standard deviations from the group mean.

4.2.2. Stimuli

The speech stimulus set was taken from Polka and Sundara (2012). A female, simultaneous bilingual speaker of Canadian French and English recorded the speech stimuli in a child-directed manner in a sound-attenuated booth. She read four passages

in each language, in which each passage consisted of six sentences. Each sentence in a passage contained a bi-syllabic target word in one of three sentence positions (beginning, middle and end). The four target words in French were *bérets, surprises, devis* and *guitares*; and the four target words in English were *candle, doctor, kingdom* and *hamlet*. The speaker was also recorded repeating each target word in isolation.

Table 4.1. Mixed dual-language passages used in the familiarization phase.

Elle a sorti ses belles **guitares**. Trois **guitares** ne seraient pas assez.
Your **hamlet** lies just over the hill. Far away from here near the sea is an old **hamlet**.
Il faut d'autres **guitares** pour la fête. Ne faites pas trop attention aux **guitares**.
People from the **hamlet** like to fish. Another **hamlet** is in the country.
On voit plusieurs **guitares** avant de choisir. Les **guitares** ne sont pas accordées.
People from that **hamlet** really like to farm. They grow so much that theirs is a very big **hamlet**.

La mode est aux **bérets** et autres chapeaux. Plusieurs **bérets** sont encore en vente.
The **candle** in the kitchen was almost melted. So Annie bought another **candle** at the stationery store.
Ces jolis **bérets** sont à ma soeur. Elle a besoin de trois **bérets**.
She came home and put away the old **candle**. Fran gave that **candle** to you later.
Il faut mettre les **bérets** sur la table. On va apporter d'autres **bérets**.
Then she made a place for the new big **candle**. Your **candle** is very pretty and smells nice too.

Voici de belles **surprises** pour vous. Il a voulu faire plusieurs **surprises**.
The **doctor** saw you the other day. He's much younger than the old **doctor**.
D'autres **surprises** risquent encore de se produire. Les **surprises** sont faciles à éviter.
I think your **doctor** is very nice. He showed another **doctor** your pretty picture.
Mieux vaut deux que trois **surprises**. Il s'attend aux **surprises** à venir.
That **doctor** thought you grew a lot. Maybe someday you'll be a big **doctor**.

Les **devis** reçus sont raisonnables. Elle a d'autres **devis** à envoyer.
Your **kingdom** is in a faraway place. The prince used to sail to that **kingdom** when he came home from school.
On doit faire confiance aux **devis**. Il y a trois **devis** posés sur la table.
One day he saw a ghost in this old **kingdom**. The **kingdom** started to worry him.
Voilà de bien beaux **devis**. Plusieurs **devis** sont falsifiés.
So he went to another **kingdom**. Now in the big **kingdom** he is happy.

Note: The above passages were French-first, English-second. Another set of passages was also used, with English-first, French-second (i.e., Sentences 1 and 2, 3 and 4, 5 and 6 in each passage were interchanged).

Three types of sound files were created for the three phases in the experiment (familiarization phase, re-familiarization phase and the test phase). For the familiarization phase, a pair of passages (one English; one French) were spliced to create four pairs of

inter-sentential, language-mixed passages. Sentences from two passages in the two languages were mixed such that the new passage switched languages every two sentences (see Table 4.1). Each of these new inter-mixed passages were repeated three times to create two-minute long passages for the familiarization phase (mean = 131.5 sec; range = 127-137 sec). Two versions of each passage were created: one with English sentences first and one with the French sentences first. Thus, a total of eight different mixed language files were created. For the re-familiarization phase, the original single-language passages were used (mean = 21.6 sec; range = 20 – 23.9 sec). For the test phase, each of the eight word lists included 12 repetitions of the target word (mean = 22.1 sec; range = 19.5 – 24.5 sec).

Each infant was familiarized to one of eight intermixed passages for 2 minutes (4 passages x 2 orders), re-familiarized with one of the four passages in each language, then tested on two of the four word lists in each language (see Table 4.2). Specifically, infants were assigned to one of eight experimental groups, which counterbalanced which passages they heard, and whether they heard the familiarization passages in English or French first. The target words were paired across languages as follows - Group I: bérets-candle, surprises-doctor, and Group II: devis-kingdom, and guitares-hamlet. For example, if infants heard the *bérets-candle* passage in the familiarization phase, they would hear the *bérets* and *candle* passage in the re-familiarization phase, and they would get tested on both the *bérets* and *candle* word lists (familiar trials) and the *surprises* and *doctor* word lists (novel trials).

Table 4.2. Experimental conditions for counterbalancing passages.

	Passive exposure	Infant-controlled HPP		
	Familiarization phase	Re-familiarization phase	Test Phase	
	2 minutes	min. of 6 sec. per passage	16 trials	
			Familiar words	Novel words
1	Guitare-Hamlet mixed passage	Guitare passage Hamlet passage	Guitare Hamlet	Devis Kingdom
2	Devis-Kingdom mixed passage	Devis passage Kingdom passage	Devis Kingdom	Guitare Hamlet
3	Bérets-Candle mixed passage	Bérets passage Candle passage	Bérets Candle	Surprises Doctor
4	Surprises-Doctor mixed passage	Surprises passage Doctor passage	Surprises Doctor	Bérets Candle

Note: The above orders are French-first, English-second. Another set of passages (i.e. English-first, French-second) was also used in the familiarization phase for a total of eight experimental conditions. In the test phase, the list of words were presented in a block-randomized order.

4.2.3. Procedure

The experiment took place in a sound-attenuated booth. There was an infant exerciser in the middle of the sound booth, and a three-sided set up of curtains around the sound booth. On the center curtain was a white light bulb, and on the side curtains were red light bulbs. Behind each side curtain was a loudspeaker.

The experiment consisted of three phases: a familiarization phase, a re-familiarization phase and a test phase. During the familiarization phase, infants were seated on the infant exerciser in the middle of the booth facing the pack, and the parent was seated in front of the infant facing the front. The parent was instructed to remain neutral, and wore headphones to avoid influencing the infant's behavior. During this

phase, we played the inter-mixed passage through the speakers, which lasted just over

two minutes. At the end of the familiarization phase, a researcher entered the sound booth

and moved the infant from the infant exerciser to the parent's lap, such that the infant was

then facing the center curtain of the three-sided set up. The researcher then removed the

infant exerciser from the sound booth.

The trials for the re-familiarization and test phase followed the same procedure.

The initiation and termination of each trial was infant-controlled. At the beginning of each

trial, the white light on the center curtain flashed. When the white light had grabbed the

infant's attention, the red light on one of the side curtains began to flash. When the infant

turned and looked to the red light, the trial was initiated and a speech file began to play

through the speaker located behind the curtain. If the infant looked at the side light until

the end of the sound file, then the white light on the center panel began flashing to start

a new trial. However, if the infant looked away from the side curtain light more than 30

degrees away for more than two seconds, the speech stimuli stopped playing and the

white light on the center panel began flashing to start a new trial. If the infant look-away

time was less than two seconds, the sound continued to play, but this look-away time was

not included in their listening time. Infants' looking was observed by a researcher via a

camera recorder connected to a computer screen outside the sound-attenuated booth.

During the re-familiarization phase, infants heard a passage play on each trial

when they looked at the light on the side curtain. The re-familiarization phase alternated

between the two language passages across trials; the trials continued until the infant

accumulated 6 seconds of listening time to each language passage, providing two

exposures to each target word in connected speech in each language while also orienting

infants to the task. The 6 second criterion for each language was typically achieved in 1-2 trials. Once the re-familiarization was reached, the test phase began.

During the test phase, the trials included the two familiar word lists (one English and one French) containing the target words that occurred throughout the familiarization and re-familiarization phase, and two novel word lists containing two words (one English and one French) that did not occur in either the familiarization or re-familiarization phase. There were 12 test trials; the two novel and two familiar word lists were presented four times in a block-randomized order. The flashing light and speech were played on the right side during half of the trials, and on the left side during the other half. The assigned side varied randomly, with the constraint that no more than three trials in a row were presented on the same side. Maximum trial length was 22 seconds; minimum trial lengths were set (3 seconds for passages; 1 second for word lists) to ensure that the infant heard the target word twice on each trial. When the minimum was not met, the trial (i.e., the same passage or word list) was immediately repeated and the short trial was removed from the analysis.

4.3. Experiment 1: Results

Prior studies have shown language-specific patterns of bi-syllabic word segmentation in young infants (e.g., Jusczyk et al., 1999). Thus, we expected that all three groups of monolingual infants would show successful segmentation in their native language, but not in their non-native language. We separated the analyses below into two types of comparisons: i) *language comparison:* 8-month-old French monolinguals vs. 8-month-old English monolinguals to examine potential linguistic effects, and ii) *age comparison:* 8-month-old French monolinguals vs. 10-month-old French monolinguals to examine potential developmental differences.

4.3.1. Language Comparison

First, we explored potential linguistic differences by examining segmentation performance in 8-month-old French and English monolinguals. Specifically, we examined whether monolingual infants were successful at segmenting bi-syllabic words in French and English. If infants listened longer to the test word than the novel word in a particular language during the test phase, then we take this as evidence that infants were able to segment the target word from the passages in that language. We averaged each infant´s listening responses to familiar and novel words to each language during the test phase, and computed the mean average listening time to familiar and novel words for each group. The group data is plotted in Figures 4.1a and 4.1b. Visual inspection of the figures suggests that the French monolinguals had a longer listening time to the French familiar word than the French novel word, but similar listening times to the English familiar and novel words. Likewise, the English monolinguals had longer listening times to the English familiar word than the English novel word, but similar listening times to the French familiar and novel words. These patterns of results are confirmed via statistical tests, as presented below.

We submitted the mean listening times scores to a mixed ANOVA with Group (French-monolingual vs. English-monolingual) as a between-subjects factor, and Nativeness (Native vs. Non-Native Language) and Trial Type (Familiar vs. Novel) as within-subjects factors. There was a main effect of Trial Type [$F(1, 30) = 19.47$, $p < .001$, $\eta_p^2 = .39$], with infants listening longer to familiar than novel trials. There was also a main effect of Nativeness [$F(1,30) = 4.83$, $p = .036$, $\eta_p^2 = .14$], with infants listening longer to trials with their native language than that of the non-native language. There was no main

effect of Group [$F(1,30) = .42$, $p = .524$, $\eta_p^2 = .01$], nor a three-way interaction between Group, Nativeness, and Trial Type [$F(1, 30) = .689$, $p = .413$, $\eta_p^2 = .02$]. Critically, as expected, there was a two-way interaction between Nativeness and Trial Type [$F(1, 30) = 14.77$, $p < .001$, $\eta_p^2 = .33$]. To understand this interaction, we conducted paired sample t-tests for each group to compare infants' listening time to the two types of trials. The French-learning infants listened significantly longer to the familiar than the novel words for the French word lists [$t(15) = 3.15$, $p = .007$, $d = .78$], but not for the English words lists [$t(15) = -1.33$, $p = .205$, $d = -.33$]. Similarly, the English-learning infants listened significantly longer to the familiar than the novel words for the English word lists [$t(15) = 3.94$, $p < .001$, $d = .99$], but not for the French word lists [$t(15) = .87$, $p = .397$ $d = .22$]. These analyses confirm that, in the test phase, the monolingual groups showed the expected familiarity effect when listening to words in their respective native language, but not when listening to words in their non-native language. Indeed, these patterns of results are robust at the individual level. The vast majority of infants exhibited a familiarity effect in their native language (i.e., 13 out of 16 French monolinguals and 15 out of 16 English monolinguals), and binomial tests indicate that these proportions are higher than expected by chance ($p = .021$ and $p = .001$, respectively). In contrast, in their non-native language, only 7 out of 16 French infants and 10 out of 16 English infants listened longer to the familiar than the novel words, and binomial tests indicate that these proportions are not significantly different from chance ($p = .804$ and $p = .454$, respectively).

4.3.2. Age Comparison

Here, we explored potential age differences by examining segmentation performance in 8-month-old and 10-month-old French monolinguals. We examined

whether French-monolingual infants would show developmental differences in segmenting bi-syllabic words in French and English. Figure 4.1c plots the group data for the 10-month-old French-monolinguals. Visual inspection of the figure shows that, like their younger 8-month-old peers, the 10-month-old French monolinguals had a longer listening time to the French familiar word than the French novel word, but similar listening times to the English familiar and novel words. To confirm this pattern, we submitted the listening times to a mixed ANOVA with Age (8-month-olds vs. 10-month-olds) as a between-subjects factor, and Nativeness (Native vs. Non-Native Language) and Trial Type (Familiar vs. Novel) as within-subjects factors. There was a main effect of Nativeness [$F(1,30) = 6.76$, $p = .014$, $\eta_p^2 = .18$], with infants listening longer to trials with their native language than that of the non-native language. As expected, there was also a two-way interaction between Nativeness and Trial Type [$F(1, 30) = 9.10$, $p = .005$, $\eta_p^2 = .23$]. Like their 8-month-old peers, the 10-month-old French monolinguals listened longer to the familiar than the novel trials in their native language [$t(15) = 2.15$, $p = .049$, $d = .53$], but not in their non-native language [$t(15) = -.74$, $p = .47$, $d = -.18$]. There was a main effect of Age [$F(1, 30) = 6.70$, $p = .015$, $\eta_p^2 = .18$], with 10-month-old infants displaying longer listening times than 8-month-old infants; critically, there was no interaction between Age, Nativeness and Trial Type [$F(1, 30) = .03$, $p = .854$, $\eta_p^2 = .001$], indicating no differences in the familiarity effect between the two age groups.
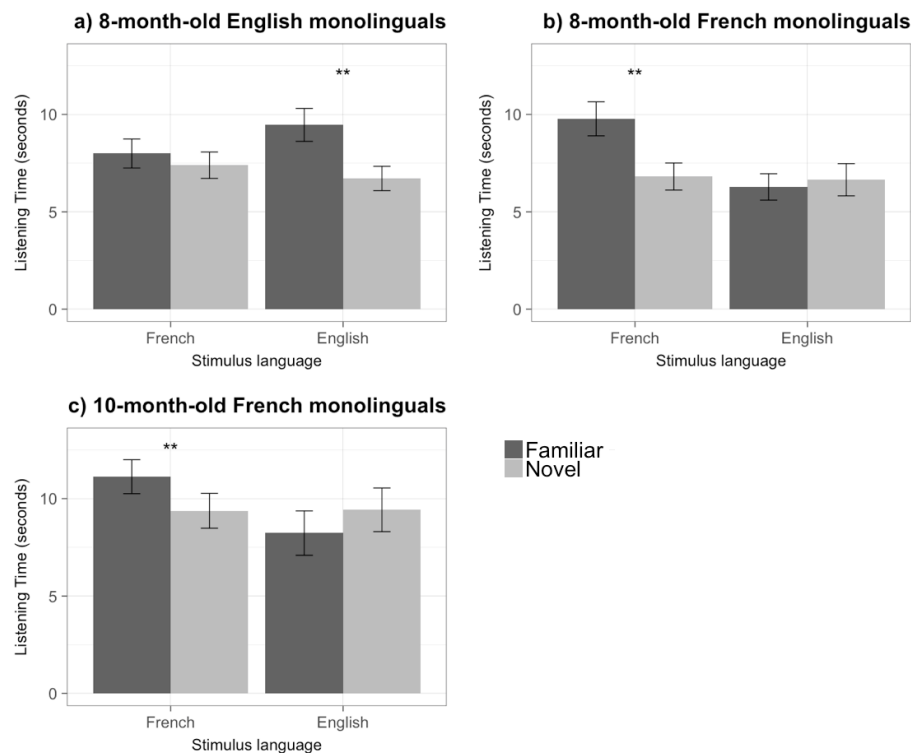
Figure 4.1. Average listening times to Familiar and Novel trials for the different group of monolingual infants: a) 8-month-old English monolinguals, b) 8-month-old French monolinguals, and c) 10-month-old French monolinguals. The error bars represent standard error. ** denotes significance (p < .05).

4.4. Experiment 1: Discussion

Our results confirm that both 8- and 10-month-old monolingual infants show language-specific patterns for word segmentation: they are able to segment bi-syllabic words in their native language, but not in their non-native language. Further, our experiment meaningfully extends previous research by showing this pattern of results in a mixed, dual-language context. In an earlier study from our laboratory, we tested the two languages in sequential blocks (Polka et al., 2017); here, the two languages were inter-mixed throughout the testing session. The finding that infants showed the same pattern of results in both types of testing conditions suggests that this pattern of results is robust.

There is no evidence here that language mixing affects word segmentation patterns for monolingual infants. The bilingual context did not bootstrap their segmentation abilities for the non-native language, nor did it disrupt their ability to segment in their native language. Instead, our findings replicate the previous findings of language-specific patterns in word segmentation.

Since all groups of monolingual infants showed the same pattern for their respective native and non-native languages, we deduced that the stimuli and task demands were comparable across the two language conditions. Specifically, we did not find any evidence that one language was more engaging than the other, nor that one language was easier for infants to segment than the other. During the test phase, there was a main effect of language for all three groups, with infants listening longer to their native language than their non-native language. However, this language preference is driven by the familiar target word in their native language, as evidenced by the interaction between language and trial type for all three groups. These results provide support for using this paradigm to examine word segmentation in 8- and 10-month-old bilingual infants.

Recall that there is sufficient evidence that bilingual infants can segment monosyllabic words in each of their native languages (e.g., Singh & Foong, 2012), but that there is limited evidence that bilingual infants can do so with more challenging bi-syllabic words. There is considerable evidence from previous studies suggesting that bilingual infants should be able to segment bi-syllabic words in each of their native languages. For example, prior studies show that bilingualism promotes learning of lexical stress patterns (Abboub, Bijeljac-Babic, Serres, & Nazzi, 2015; Bijeljac-Babic, Serres,

Höhle, & Nazzi, 2012), which in turn can facilitate word segmentation. In addition, some studies show that the bilingual experience has minimal impact on children's statistical learning abilities (e.g., Yim & Rudoy, 2013), suggesting that bilingual infants could also use learned speech cues at the same pace as monolingual infants. Indeed, prior research shows that bilingual infants can use different prosodic or indexical cues to segment phrases from an artificial language (Gervan & Werker, 2013; Tsui, Erickson, Thiessen & Fennell, 2017).

Further, a recent study by Antovich and Graf Estes (2017) examined early word segmentation in a bilingual context using artificial languages. They found that while 14-month-old monolingual infants could only segment bi-syllabic words in one of two language streams, their bilingual peers were able to segment words from both language streams. These results suggest that early bilingual exposure supports the ability to separate language streams and extract bi-syllabic words from those streams. If these abilities can scale up to natural language processing contexts and are available earlier in life, then young bilingual infants should show successful segmentation in both of their native languages in our task.

In Experiment 2, we tested the word segmentation abilities of infants growing up in French-English bilingual homes. The data from our previous study showed that 8-month-old infants had difficulty segmenting in both of their native languages when the languages were tested in successive blocks. In Experiment 2, we examined the segmentation abilities of both 8- and 10-month-old bilingual infants in this new mixed dual-language segmentation task.

4.5. Experiment 2: Methods

4.5.1. Participants

We recruited two groups of infants from French-English bilingual homes from Montréal, Canada: eight-month-old infants (n=16, Mean age = 238.0 days; Age range = 224 – 259 days; 9 boys), and ten-month-old infants (n=20, Mean age = 302.8 days; Age range = 289 – 325 days; 9 boys). Parent-reported measures indicated that infants received at least 30% exposure to each of their ambient languages, and no exposure to any other languages. An additional seven infants were tested but excluded from the final sample because of fussiness (2), language background criteria exclusion (3) and technical error (2). Further, three infants were identified and excluded as outliers using the same pre-established criteria as experiment 1.

During their visit to the laboratory, we interviewed parents about their child's language background at home. Preliminary data show that parent-reports about their child's language exposure are highly reliable, when compared against transcriptions of home recordings using the Language Environment Analysis (LENA) recording system (Orena, Byers-Heinlein & Polka, in preparation). For both age groups, there was an equal number of French- and English-dominant infants, with dominance defined as the language the child was most exposed to based on parental reports. The primary caregiver also completed the Language Mixing Scale (LMS; Byers-Heinlein, 2012), which asks parents to report how much they engage in intre-sentitial code switching between their two languages to their child (possible scores = 0 – 30, with higher scores indicating more mixing). For our bilingual infants (collapsed across age groups) the mean LMS score was 12.0 (SD = 8.04; Range = 0 – 30). We had incomplete data for three infants; these

subjects were excluded from any analyses of segmentation performance and LMS scores.

4.5.2. Stimuli

Same as in Experiment 1.

4.5.3. Procedure

Same as in Experiment 1.

4.6. Experiment 2: Results

First, we examined whether bilingual infants were successful at segmenting bi-syllabic words in their dominant and non-dominant languages in this mixed dual-language task. We conducted our analyses by dominance (i.e., Dominant vs. non-dominant language) instead of by language (i.e., French vs. English) in order to examine whether bilingual infants would show stronger patterns of segmentation in their more "familiar" language, as is the pattern with monolingual infants in Experiment 1[3]. We computed infants' listening responses to familiar and novel words in each language during the test phase, and plotted the group data in Figures 4.2a and 4.2b. Visual inspection of the figures suggests that both groups of bilingual infants listened longer to the familiar words than the novel words in both their dominant and non-dominant languages.

We submitted these means to an ANOVA with Age (8-month-olds vs. 10-month-olds) as a between-subjects factor, and Dominance (Dominant vs. Non-Dominant

---

[3] When we analyzed the data in an ANOVA with language (French vs. English) the same pattern was observed, we found that, as a group, both 8- and 10-month-old infants were segmenting in both languages. In an ANOVA with age as a between-subjects factor, and language and trial type as within-subjects factors, there was a main effect of trial type ($p < .001$), but no other main effects or interaction (all $ps > .250$).

Language) and Trial Type (Familiar vs. Novel) as within-subjects factors. There was a main effect of Trial Type [$F(1,34) = 17.44$, $p < .001$, $\eta_p^2 = .34$], indicating that both groups of infants listened longer to the familiar words than the novel words in both languages. There were no other significant main effects or interactions, including no interaction between Dominance and Trial Type [$F(1,34) = .42$, $p = .522$, $\eta_p^2 = .01$], and no interaction between these two factors and Age [$F(1,34) = .16$, $p = .689$, $\eta_p^2 = .005$]. These results suggest that both 8- and 10-month-old bilingual infants are able to segment in both their dominant and non-dominant languages. Indeed, planned paired t-tests show that 8-month-old bilingual infants listened significantly longer to the familiar word than the novel word in their dominant language [$t(15) = 5.57$, $p < .001$, $d = 1.39$], and marginally significantly so in their non-dominant language [$t(15) = 2.01$, $p = .063$, $d = .50$]. Similarly, 10-month-old bilingual infants listened significantly longer to the familiar than the novel word in their dominant [$t(19) = 2.07$, $p = .052$, $d = .46$] and non-dominant [$t(19) = 2.58$, $p = .018$, $d = .58$] languages.

However, note that compared to the monolinguals, the mean listening ratio and effect sizes of the listening time analysis for the bilingual group are numerically smaller than that of the monolingual groups. One possibility is that there may be more individual variability in the bilingual group, with some infants showing a familiarity effect in one language but not the other. For words in their dominant language, 27 out of 36 bilingual infants listened longer to the familiar word than the novel word, and a binomial test indicates that these proportions are higher than expected by chance ($p = .004$). Similarly, for words in their non-dominant language, 25 out of 36 bilingual infants listened longer to the familiar word than the novel word, and these proportions are also higher than

expected by chance (p = .029). However, only 20 out of 36 infants listened longer to the familiar word than the novel word in both languages, and this proportion is not significantly different by chance (p = .618). Indeed, visual inspection of the data suggests that there is greater individual variability in how bilingual infants performed in this task.
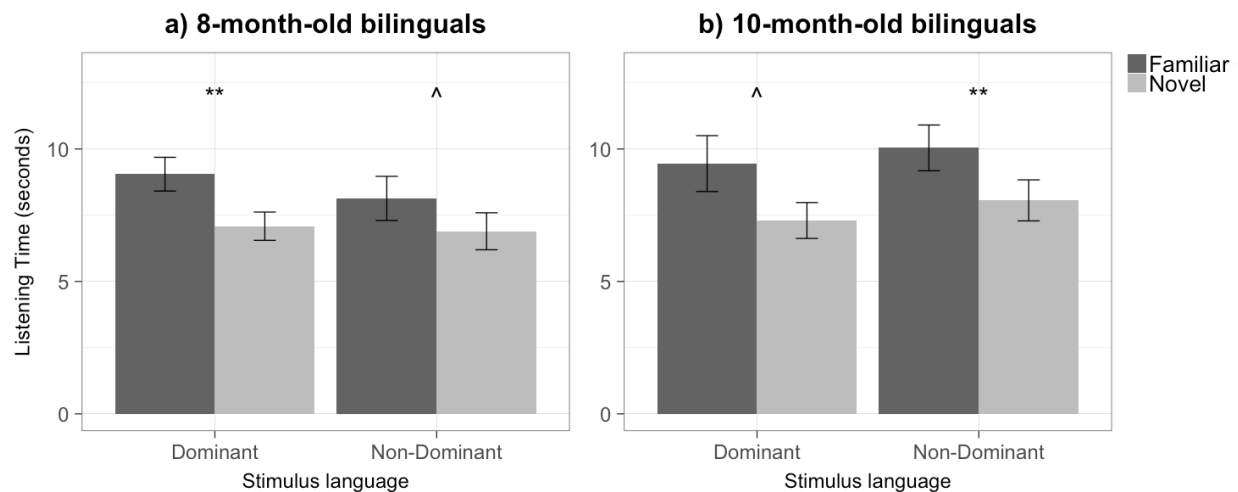
Figure 4.2. Average listening times to Familiar and Novel trials for the different age group of bilingual infants: a) 8-month-olds, and b) 10-month-olds. The error bars represent standard error. ** denotes significance (p < .05), and ^ denotes marginal significance (.05 < p < .07).

One possibility is that the amount of language mixing that infants hear may mediate performance in this task, such that bilingual infants who are exposed to more language mixing may have greater success in segmenting both languages in this task compared to those who have less experience with mixing. Exploratory analyses suggest that this may be the case. First, we divided infants into two groups based on a median split of the scores in their parent-reported language mixing scale; three participants had incomplete data and were excluded from this analysis. There were 17 participants in the high mixed input group, with an LMS mean of 18.5 (SD = 5.08; Range = 11 - 30), and 16 participants in

the low mixed input group, with an LMS mean of 5.06 (SD = 3.26; Range = 0 - 10). The group data are plotted in Figures 4.3a and 4.3b. We submitted infants' listening times during the test trials to a 3-way ANOVA with Mixing (low mixed input vs. high mixed input) as a between-subjects factor, and Dominance and Trial Type as within-subjects factors, which found an interaction between the three factors [$F(1,31) = 6.26$, $p = .018$, $\eta_p^2 = .17$]. Indeed, infants who experienced more language-mixing from their mothers listened longer to the familiar than the novel word in their dominant language, albeit with only marginal significance [$t(16) = 1.85$, $p = .08$, $d = .45$], as well as in their non-dominant language [$t(16) = 3.15$, $p = .006$, $d = .77$]. In contrast, infants who experienced less language-mixing from their mothers showed a pattern similar to that of monolingual infants; their segmentation performance was clearly robust in their dominant language [$t(15) = 3.16$, $p = .007$, $d = .84$], but not in their non-dominant language [$t(15) = 1.28$, $p = .220$, $d = .35$].
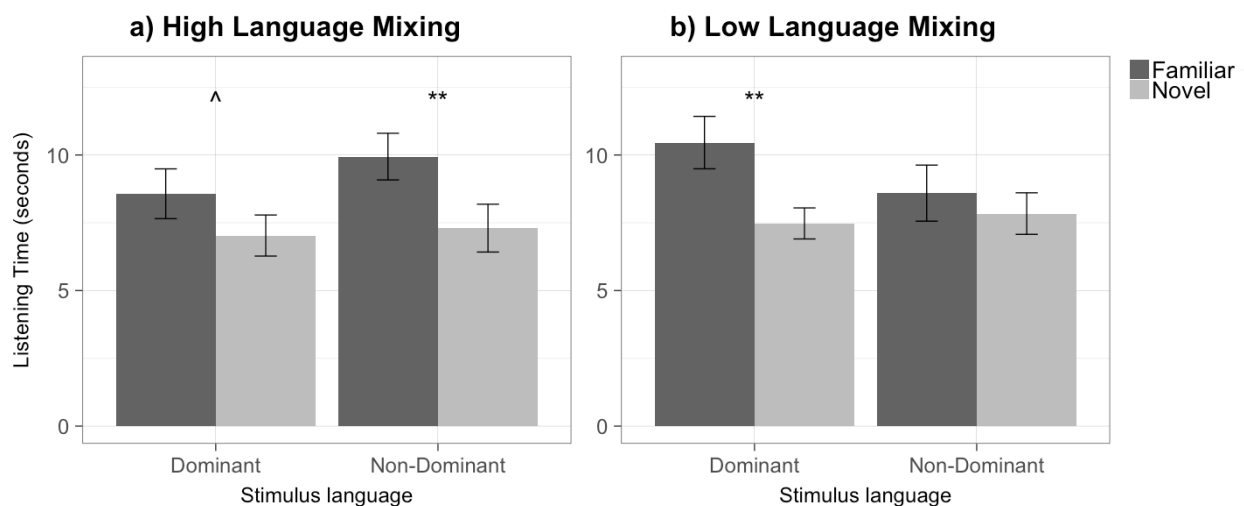

Figure 4.3. Average listening times to Familiar and Novel trials for the group of bilingual infants divided by amount of exposure to language mixing: a) High amounts of language mixing, and b) Low amounts of language mixing. The error bars represent standard error. ** denotes significance ($p < .05$), and ^ denotes marginal significance (.05 < p < .08).

Indeed, these findings hold when treating the language mixing scale as a continuous variable, with infants' segmentation ratio as the dependent variable (i.e., listening time during familiar trials divided by listening time to novel plus familiar trials). In this analysis, a segmentation ratio of 0.5 represents equal listening time to familiar and novel words, while a segmentation ratio above 0.5 represents a preference for the familiar word over the novel word. Here, we find no significant correlation between a mom's score on the language mixing scale and their infant's segmentation index for their dominant language [r(31) = -.20, p = .28] (see Figure 4.4). However, there was a significant moderate correlation between these two variables for the infants' non-dominant language [r(31) = .39, p = .03]. This analysis suggests that bilingual infants are able to segment in their dominant language, regardless of the amount of language mixing they receive from their caregivers; critically, those who receive more language mixing from their caregivers are also able to segment words in their non-dominant language. The pattern of data from these exploratory analyses suggests that the types of speech input bilingual infants receive may play a role in their performance in word segmentation tasks.
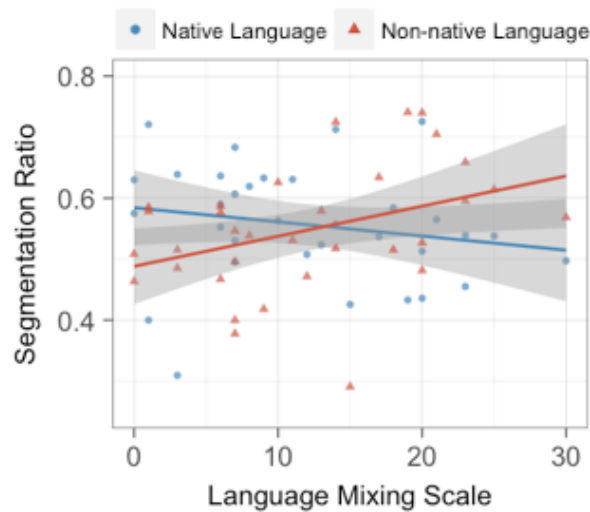
Figure 4.4. The relationship between exposure to language mixing and infants' word segmentation skills in their native language (blue circles) and their non-native language (red triangles).

4.7. Experiment 2: Discussion

Our results show that, when tested with natural connected speech, bilingual infants successfully segment in each of their native languages by 8-months of age. Bilingual infants from both age groups show stable segmentation of their more dominant language, and even by conservative standards, their ability to segment in their less dominant language emerges at 8-months of age and stabilizes by 10-months of age. This is the first direct evidence that the same group of bilingual infants can segment bi-syllabic words in each of their native languages by 8-months of age.

These results are striking given the challenges of dual-language word segmentation. First, segmenting bi-syllabic words from two languages that differ in rhythmic structure is no easy task. Indeed, while young monolingual infants can segment words that follow the same prosodic structure from a non-native language stream (Houston et al., 2000; Lew-Williams, Pelucchi, & Saffran, 2011), they are unable to

segment words that have prosodic structures that are atypical to them until 12-months of age (Jusczyk et al., 1999). We can conclude from this study that given sufficient language experience, infants can learn to segment words with different prosodic structure by 8-months of age. Second, the French and English stimuli that we used in this task were spoken by the same speaker; thus, infants did not have access to indexical cues for separating the language streams. Nevertheless, bilingual infants were successful in segmenting words from the two language streams – implying either that they have learned a universal template for segmenting in both languages, or that they were able to sort the two language streams and use the appropriate strategy for segmenting in each stream. Regardless of the explanation, bilingual infants accomplish the impressive feat of dealing with variability in their speech input and breaking into two language systems.

One noteworthy aspect of these results is that there appears to be wider variability in performance for the bilingual infants compared to the monolingual infants. Indeed, our exploratory analyses showed that language mixing may affect word segmentation performance in this task. It is not surprising that infants who hear a lot of language mixing in their input would do better at a word segmentation task in a language-mixed task. However, other factors may also be in play here. Our exploratory analyses show a relation between the two variables, but there may be other language experiences that may mediate this relation. Certainly, bilingual groups tend to be more heterogeneous than monolinguals in terms of language experiences. For example, bilingual families can differ in terms of the caregiver's nativeness and fluency in both of their languages, the amount of talk in each of their languages, and the contexts or activities in which infants hear these languages – and each of these factors could affect segmentation processes for infants.

For instance, it is common for bilingual infants to be exposed to foreign-accented speech in at least one of their languages from one or both of their caregivers, which in turn may cause infants to incorrectly learn a cue for segmentation. For example, a French-English caregiver might produce English content words with a French accent (e.g., producing the word *TA-ble* as an iambic *ta-BLE,* the structure more common in French). If infants learned such a pattern, they would be able to segment words in speech from their parents, but not necessarily in speech from the community (as is the case with our stimuli). The variety of language experiences in our bilingual group could be the source of the noise in the data. These analyses suggest that exploring the types of input bilingual infants receive would be a promising direction for future research.

4.8. General Discussion

The goal of the current investigation was to examine the nature of infants' word segmentation in bilingual contexts. Several studies have shown that the ability to segment words is facilitative for word learning and vocabulary development (Estes, Evans, Alibali, & Saffran, 2007; Newman, Rowe, & Bernstein Ratner, 2015; Singh, Reznick, & Xuehua, 2012). Thus, it is important to understand the external factors that contribute to this challenging skill. Here, we explore the relation between language exposure and word segmentation. Our findings support three conclusions. First, broad language familiarity bolsters word segmentation skills. Second, bilingual infants possess the processing capacity to segment bi-syllabic words from two rhythmically-different languages, and this skill emerges when they are provided sufficient language experience. Third, different types of bilingual experience may affect the precise trajectory of the infants' speech processing abilities in each language. We discuss the data that support these conclusions

below, and their implications for our understanding of speech processing and bilingual language development.

Word segmentation as a language-specific skill

As expected, monolingual infants showed language-specific patterns of word segmentation: they showed successful segmentation in their native language, but not in a non-native language. It is noteworthy to add that this pattern held, even in the face of a bilingual stream. Regardless of whether monolingual infants were tested in two languages sequentially (as in Polka et al., 2017), or in two languages in a language mixed task (as in this study), the bilingual context did not appear to affect this general language-specific pattern. Similarly, bilingual infants showed language-specific patterns of word segmentation in that they are able to segment words in both of their native languages by 8-months of age. These results are consistent with the hypothesis that language exposure is a critical factor for one's ability to segment words in that language.

There are at least two non-mutually exclusive routes for this phenomenon. The first, as alluded to in the introduction of this paper, follows the prosodic bootstrapping hypothesis (Morgan & Demuth, 1996), which suggests that young language learners can break into the structure of their native language using prosodic features (such as pitch and rhythm). Research with adults confirm that those who speak a stress-timed language (e.g., English, German) are biased to track and segment words using stress, while those who speak a syllable-timed language (e.g., French, Spanish) use syllables to track units from the speech stream (e.g., Cutler et al., 1986). Certainly, prior research on early word segmentation confirms that young infant can segment bi-syllabic words that follow their native language word stress pattern, even when they are tested in a rhythmically similar

but unfamiliar language or dialect (Houston et al., 2000; Polka & Sundara, 2012). The same explanation can be extended to bilingual infants. Bilingual infants appear to follow the same developmental path in learning to process word stress cues as their monolingual peers, even when one of their languages is not stress-timed (French-German bilinguals: Bijeljac-Babic, Höhle, & Nazzi, 2016). These prior studies support the hypothesis that the French-English bilinguals in this study were able to learn prosodic features of their native languages, and apply them for segmentation in each respective language.

Alternatively, our pattern of results could also be explained by attention mechanisms. In our task, two languages are being presented within the same discourse. It is possible that infants are only paying attention when the incoming speech matches what they are regularly exposed to in their daily lives. Under the PRIMIR framework, Werker and Curtin (2005) posit that infants prioritize certain information when processing speech, especially in processing-heavy contexts. Certainly, from birth, monolingual infants prefer listening to their native language, while bilingual infants show equal preference for both of their native languages (Byers-Heinlein, Burns, & Werker, 2010). Further, Marno and colleagues (2016) show that infants prefer learning from speakers that speak their native language. Thus, it is quite possible that monolingual infants only segmented in their native language because they devoted more attention to the speech stream when it conformed to their phonological, prosodic and phonotactic expectations. Similarly, it is possible that bilingual infants, and not monolingual infants, showed segmentation to both language streams because both languages are equally salient to them. Future research is needed to explore this alternative hypothesis.

Bilingual infants' processing capacity

The current data shows that bilingual infants have the processing capacity to segment bi-syllabic words from two different languages, and that this ability depends on language experience: bilingual infants, but not monolingual infants, are able to segment words from two language streams. These results are striking given the challenges of dual-language word segmentation. Nevertheless, research with artificial languages show that bilingual infants (but not monolingual infants) can track prosodic and syllabic cues from two streams of speech (Antovich & Graf Estes, 2017; Gervain & Werker, 2013), suggesting that bilingual infants have the capacity to learn and use prosodic cues in a context-specific manner. Our current study shows the same type of processing with a dual language stream of English and French, indicating that this skill holds validity in a natural language context. Further, in the studies that use artificial languages, indexical information is typically used to separate the two language streams (Antovich & Graf Estes, 2017; Gervain & Werker, 2013). Our study suggests that infants do not need indexical information to separate the two language streams for word segmentation purposes.

How do we reconcile our current results with Polka and colleagues (2017), which showed that bilingual infants have difficulty segmenting in both of their native languages? There are several methodological differences between these two studies which may explain the discrepancy of results. First, Polka and colleagues (2017) tested the languages in sequential blocks, while the study here tested them in a language-mixed task. One possibility is that infants are better able to process speech in the way that they experience it at home (Mattock, Polka, Rvachew, & Krehm, 2010). The group LMS score

reveals that some parents do a fair amount of language mixing when they are talking to their children. As such, it is possible that bilingual infants succeeded in segmenting in this language-mixed task because this type of language mixing is more common to them at home.

It is interesting to note that monolinguals had no trouble segmenting in their native language in both studies, suggesting that the monolinguals were unfazed by either bilingual context. Indeed, one of the extra challenges that bilingual face when encountering a new speaker is resolving the initial ambiguity of what languages their interlocutors can speak. Thus, we can conclude that while bilingual infants do have the processing capacity to segment in both of their languages, this ability may require more perceptual adaptation from the bilingual infants to cope with differing language modes.

Finally, the language exposure during the familiarization phase differed slightly. In Polka and colleagues (2017), the familiarization phase was infant-controlled and provided forty-five seconds of exposure to each passage. In the present study, infants were provided two minutes of passive exposure to both languages, along with at least 6 seconds of exposure to each language controlled by infant looking behavior. We hypothesize that this longer exposure to the passages helped infants by giving them more time to segment words from the passage. Indeed, Nazzi and colleagues (2014) found that even minimally different durations of exposure (30 seconds vs. 45 seconds) could be the difference between showing the ability to segment and not in such a task.

Type of bilingual experience may affect word segmentation

Interpreting findings from studies with a bilingual population is often complex due to the wide heterogeneity in experiences among bilingual families. Indeed, there is limited

research about how specific parameters of the bilingual experience shape infants' language abilities. Certainly, the linguistic landscape within bilingual families can differ in many ways. For example, bilingual families can differ in the pair of languages that they speak, the caregiver's nativeness and fluency in both of their languages, the amount of talk infants hear in each of their languages, and the contexts or activities in which infants hear these languages. Further, some families might adopt a one-person-one-language structure of language exposure for their infants, while others freely switch between their two languages. All of these variable factors make it challenging to identify the critical components of the bilingual experience that affect their language abilities.

In our exploratory analyses, we examined how language mixing might affect word segmentation abilities in bilingual infants. Our task presents the sentences in a language-mixed discourse, so the prediction would follow that those who hear more language mixing at home would be better able to process the language-mixed discourse in our experiment. Indeed, that is what we found. In our bilingual sample, some parents often mixed their languages with their infants, and it is those infants that showed more stable performance in segmenting in both of their native languages. In contrast, those who did not hear a lot of language mixing from their parents showed successful segmentation in their dominant language, but not their non-dominant language. These results suggest that some bilingual infants can keep track of the languages despite the language mixing in their environment, and can process the languages in an efficient manner. A recent study by Byers-Heinlein and colleagues (2017) supports this hypothesis by showing that by their second birthday, infants can keep track of language switches, even when the mixing happens within the same sentences. Nevertheless, caution should be taken in interpreting

our language mixing data, especially given its exploratory nature. It is possible that other confounding factors mediate the relation between language mixing and segmentation performance, including their caregivers' language proficiency. Nevertheless, these results suggest that the wider variability in performance by bilingual infants can be explained by more detailed analyses of infants' language experiences.

Conclusions

In this paper, we examined how infants tackle the challenging task of word segmentation when the input stream contains two languages. Our results show that 8- and 10-month-old monolingual infants persist in segmenting in their native language but not in a non-native language, while bilingual infants show segmentation for both of their native languages. These findings show that infants are capable of segmenting words from natural streams of speech in two languages that are rhythmically different when given sufficient language experience. Future research is needed to elucidate how variation in bilingual language exposure modulates the trajectory of early language processing skills in bilingual language learners.

4.9. References

Abboub, N., Bijeljac-Babic, R., Serres, J., & Nazzi, T. (2015). On the importance of being

bilingual: word stress processing in a context of segmental variability. *Journal of Experminetal Child Psychology*, *132*, 111–120. https://doi.org/10.1016/j.jecp.2014.12.004

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

Antovich, D. M., & Graf Estes, K. (2017). Learning across languages: bilingual experience supports dual language statistical word segmentation. *Developmental Science*, e12548. https://doi.org/10.1111/desc.12548

Bernard, A., & Onishi, K. H. (submitted). Novel phonotactic learning by children and infants: Generalizing syllable-position but not co-occurrence regularities.

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 of the United States of America*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bergmann, C., & Cristia, A. (2015). Development of infants' segmentation of words from native speech: a meta-analytic approach. *Developmental Science, 19*(5), 901-917. https://doi.org/10.1111/desc.12341

Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2016). Early Prosodic Acquisition in Bilingual Infants: The Case of the Perceptual Trochaic Bias. *Frontiers in Psychology*, *7*, 210. https://doi.org/10.3389/fpsyg.2016.00210

Bijeljac-Babic, R., Serres, J., Höhle, B., & Nazzi, T. (2012). Effect of bilingualism on lexical

stress pattern discrimination in French-learning infants. *PLoS One*, *7*(2), e30843. https://doi.org/10.1371/journal.pone.0030843

Bosch, L., Figueras, M., Teixidó, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, *4*, 106. https://doi.org/10.3389/fpsyg.2013.00106

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2–3), 217–243. https://doi.org/10.1177/00238309030460020801

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

Byers-Heinlein, K. (2012). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism*, *16*(1), 32–48. https://doi.org/10.1017/s1366728912000120

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

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

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the

English vocabulary. *Computer Speech & Language*, *2*(3), 133–142. https://doi.org/10.1016/0885-2308(87)90004-0

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*(4), 385–400. https://doi.org/10.1016/0749-596X(86)90033-1

Dahan, D., & Bernard, J. M. (1996). Interspeaker variability in emphatic accent production in French. *Language and Speech*, *39*(4), 341–374. https://doi.org/10.1177/002383099603900402

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress deafness. *The Journal of the Acoustical Society of America*, *110*(3), 1606–1618. https://doi.org/10.1121/1.1380437

Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., White, L., Goslin, J., & Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition*, *148*, 1–9. https://doi.org/10.1016/j.cognition.2015.12.004

Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, *110*, 1–17. https://doi.org/10.1016/j.ijpsycho.2016.10.004

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, *4*, 1490. https://doi.org/10.1038/ncomms2430

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

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262–274. https://doi.org/10.1016/j.infbeh.2009.03.004

Hohle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127. https://doi.org/10.1111/1467-7687.00261

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, *7*(3), 504–509. https://doi.org/10.3758/BF03214363

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5,* 69-95.

Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental Science*, *12*(1), 131–141. https://doi.org/10.1111/j.1467-7687.2008.00740.x

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*(9), 323-328. https://doi.org/10.1016/S1364-6613(99)01363-7

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant

stress patterns of English words. *Child Development*, *64*(3), 675–687. https://doi.org/
10.2307/1131210

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3), 159–207. https://doi.org/10.1006/cogp.1999.0716

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, *14*(6), 1323–1329. https://doi.org/10.1111/j.1467-7687.2011.01079.x

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, *6*(5), 314–317. https://doi.org/
10.1111/j.1467-9280.1995.tb00517.x

Marno, H., Guellai, B., Vidal, Y., Franzoi, J., Nespor, M., & Mehler, J. (2016). Infants' Selectively Pay Attention to the Information They Receive from a Native Speaker of Their Language. *Frontiers in Psychology*, *7*, 1150. https://doi.org/10.3389/fpsyg.2016.01150

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

Morgan, J. L., & Demuth, K. (1996). Signal to syntax: An overview. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 1–22.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for

crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283–299. https://doi.org/10.1016/j.jml.2005.10.004

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language*, *41*(3), 600–633. https://doi.org/10.1017/S0305000913000111

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2015). Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, 1–16. https://doi.org/10.1017/S0305000915000446

Orena, A. J., Byers-Heinlein, K., & Polka, L. (in preparation). Reliability of the Language Environment Analysis (LENA) in French-English Bilingual Speech.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants. *Child Development*, *80*(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Polka, L., & Sundara, M. (2012). Word Segmentation in Monolingual Infants Acquiring Canadian English and Canadian French: Native Language, Cross-Dialect, and Cross-Language Comparisons. *Infancy*, *17*(2), 198–232. https://doi.org/10.1111/j.1532-7078.2011.00075.x

Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2017). Segmenting words from fluent speech during infancy – challenges and opportunities in a bilingual context. *Developmental Science*, *20*(1). https://doi.org/10.1111/desc.12419

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old

Infants. *Science,* *274*(5294), 1926-1928. https://doi.org/10.1126/science.274.5294.1926

Sandoval, M., & Gómez, R. L. (2016). Overriding the Metrical Bias with Lexical Information: English-Learning 7.5-Month-Olds UseMommyto Segment Iambic Words. *Language Learning and Development*, *12*(4), 398–412. https://doi.org/10.1080/15475441.2016.1162722

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3. https://doi.org/10.1037/0012-1649.33.1.3

Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*(2), 128–142. https://doi.org/10.1016/j.cognition.2012.05.008

Singh, L., Reznick, S. J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, *15*(4), 482–495. https://doi.org/10.1111/j.1467-7687.2012.01141.x

Statistics Canada (2016). Focus on Geography Series, 2016 Census. *Statistics Canada Catalogue no. 98-404-X2016001*. Ottawa, Ontario.

Syrnyk, C., & Meints, K. (2016). Bye-bye mummy - Word comprehension in 9-month-old infants. *British Journal of Developmental Psychology,* *35*(2), 202-217. https://doi.org/10.1111/bjdp.12157

Tsui, A., Erickson, L. C., Thiessen, E., & Fennell, C. (2017). Statistical Learning from Accented Speech: A Bilingual Advantage. *Proceedeings of the 41st Annual Boston University Conference on Language Development, 679 - 690.* Sommerville, MA:

Cascadilla Press.

Van de Weijer, J. (1999). *Language Input for Word Discovery*. Radboud University, Nijmegen.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234. https://doi.org/10.1080/15475441.2005.9684216

Yim, D., & Rudoy, J. (2013). Implicit Statistical Learning and Language Skills in Bilingual Children. *Journal of Speech Language and Hearing Research*, *56*(1), 310. https://doi.org/10.1044/1092-4388(2012/11-0243)

CHAPTER 5. General Discussion

In this dissertation, we examined the nature of bilingualism in infancy. The results from this dissertation contribute to a growing body of literature demonstrating that infants have a wide variety of language experiences (e.g., Tamis-Lemonda, Custode, Kuchirko, Escobar, & Lo, 2018; Tamis-Lemonda, Kuchirko, & Song, 2014), and that these experiences matter for speech processing and language skills (e.g., Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Marchman et al., 2016; Ramírez-Esparza, García-Sierra, & Kuhl, 2016). Below, we briefly summarize the results of each chapter.

5.1. Summary of Results

The central goal of this dissertation was to examine the role of the language input in bilingual infants' language skills. Thus, the first step of this work was to describe how bilingual infants experience this language input. To do so, we made use of the Language ENvironment Analysis (LENA) recording system to build an audio corpus that documents what bilingual infants hear in a typical day at home. One of the features of this LENA recording system is that it includes software that estimates the number of words spoken in the recording by an adult – a number that reflects the quantity of language input to the child. The algorithms that perform this estimation were designed for counting English words; thus, in *Chapter 2,* we examined how the LENA system would also count words for both Canadian English and Canadian French. Indeed, assessing the reliability of the LENA algorithms is important for validating its use for a French-English population in both a research or clinical context. Here, we found that the LENA algorithms were equally reliable at estimating the number of words in both Canadian English and Canadian French. Further, we found that this accuracy does not deteriorate when looking at

potential differences between estimating words from male and female voices, or from accented and non-accented speech. This study adds to the growing number of papers that have validated LENA-generated estimates in different languages. Thus, this short paper provides support for using the LENA system for obtaining measures of input quantity in a bilingual population.

Another way to assess a child's language environment is to conduct language interviews with the child's caregivers. Indeed, this traditional method is preferred in both research and clinical contexts for its practicality. However, despite its widespread use, very little is known about how accurate caregivers are at describing their child's language environment. Indeed, especially in highly bilingual homes, one could imagine the difficulty of tracking what languages are being spoken and by whom. In *Chapter 3,* we provide evidence that these caregiver reports are sufficiently reliable. Specifically, we found that measures of language exposure (i.e., proportion of exposure to each language) collected during the language interview were highly correlated with measures of language exposure from the naturalistic recordings collected via the LENA system. Further, we also provide evidence that parents are reliable at describing their own language use to their child, as well as their own language policy strategies for providing bilingual input to their child. We have also presented other data analysis, not included in this chapter, that caregiver responses to the Language Mixing Scale (parent-report measure of language mixing to the infant; Byers-Heinlein, 2012) was also correlated with the amount of language mixing found in the LENA recordings (Orena, Srouji, Byers-Heinlein & Polka, 2018). Thus, this paper provides support for using language interviews for assessing a bilingual child's language experiences.

In the same chapter (*Chapter 3)*, we also touched upon the variability of language experiences amongst our group of French-English bilingual infants. This group of bilingual infants is representative of children whom we might recruit in a typical study in Montréal; thus, it was important to examine the variability of bilingual experiences in a cohort of infants that would typically be considered representative of "bilingual infants" in a group-level comparison. Here, we found natural variability in amount of language exposure to each language, in proportional exposure to each language, amount of exposure to language mixing, as well as caregiver fluency in each language. Further, we found that while most caregivers were consistent in their language use from day-to-day (at least at 10-months of age), they were not always consistent in their language use depending on who they were speaking to. These findings add to the growing literature showing individual differences in language experiences amongst not only bilingual infants, but also monolingual infants (Tamis-Lemonda et al., 2018).

Finally, in *Chapter 4*, we examined how the bilingual experience might affect speech processing. Specifically, we examined whether monolingual and bilingual infants could segment bisyllabic words from a natural speech stream with two different languages. This difficult task examines infants' processing capacities for segmenting word forms in a mixed-language stream, as well as their ability to use language-specific knowledge for word segmentation. Here, we found that 8- and 10-month-old monolingual infants could only segment words from their own language, reflecting the language-specificity of the task. Further, the 8- and 10-month-old bilingual infants showed evidence of segmentation in both of their native languages. Exploratory analyses indicate that infants who hear more language-mixing from their caregivers were the ones who showed

dual language segmentation in this task, while those who heard less language-mixing from their caregivers only segmented words from their dominant language. These findings reveal the processing capacities of bilingual infants for word segmentation, and is supported by previous computational work (Fibla Reixachs, 2017), as well as behavioral work with artificial languages (Antovich & Graf Estes, 2017).

5.2. Take-home messages

While the three manuscripts in this dissertation each had different research motivations, they all addressed the nature of bilingualism in infancy. Below, we outline two collective take-home messages from this dissertation, and discuss their implications for research on bilingual development.

- **Conclusion 1:** <u>Assessing research methods is a worthwhile venture for broadening our understanding of language acquisition.</u>

In this dissertation, we examined the utility of using caregiver interviews and the LENA system for assessing language input in bilingual infants. These were necessary methodological steps for justifying current approaches of examining infants' language experiences. Indeed, to ensure the validity of our interpretation of our results, we must consider how the variables that we are measuring operationalize the concepts that we are actually evaluating.

Firstly, from a research perspective, validating the use of caregiver interviews allows future researchers to be confident about its operationalization. While there will always be measurement errors or biases with caregiver reports, our study shows that these differences are minimal. As we continue to use these caregiver reports, we must

consider both biases from the interviewer and caregivers so as to minimize these errors (Byers-Heinlein et al., submitted).

Second, validating the LENA system opens up the potential for investigating the input quantity in more detail. For monolingual infants, this absolute measure of input quantity appears to be the critical variable for some language outcomes (e.g., Huttenlocher et al., 1991). However, for bilingual infants, most researchers have only examined the effects of the proportional measures of input quantity (collected via language interviews). As discussed in Chapter 3, this proportional variable may not tell the complete story of the input quantity. Thus, the LENA recording system allows for the opportunity to obtain absolute measures of input quantity, and when used in conjunction with other speech processing or language tasks, it can provide us the opportunity to examine how the absolute and proportional measures of bilingual input interact in predicting language outcomes.

From a clinical perspective, these two tools may also be important for assessing an infant's language experiences. Consider, for example, how clinicians might assess a bilingual child's language skills. Several researchers suggest taking children's experiences with each language into account, especially when comparing against monolingual norms (Gathercole, Thomas, Roberts, Hughes & Hughes, 2013; Thordardottir, 2015). Indeed, collecting this information is important to determine the source of language problems for a bilingual child who scores lower than the monolingual norms. Several clinical assessments that make use of caregiver reports have been developed to assess a bilingual child's language experiences and language skills (Paradis, Emmerzael, & Duncan, 2010; Tuller, 2015), and our study provides support that

caregiver reports can be a valid way to estimate a child's language experiences (*Chapter 3*). Further, some educational programs are using the LENA system to improve parent-child interactions and children's language skills (see [https://lena.org/solutions/)](https://lena.org/solutions/), and our study provides support for using the LENA system in a bilingual population (*Chapter 2).*

- **<u>Conclusion 2:</u>** <u>Infants' language experiences vary widely, and these differences matter for their language development.</u>

There are many different circumstances that can lead to a person being bilingual. In our studies, we focused on infants who simultaneously learn two languages from birth (also known as *dual-language learners*). Yet, even within our sample of French-English bilinguals, there was wide variability in the amount of language input they received, from whom, and in what contexts. One could only imagine the additional variability in language experiences when including infants learning other languages, infants who are hearing more than two languages at home, infants who receive exposure to their second language in other contexts (e.g., daycare), and infants who do not receive exposure to their second language until later in life. It is important to consider these input factors, especially when interpreting results from group-level comparisons. While bilinguals have the common characteristic of being able to speak two languages, the manner in which they learned their two languages and the way they represent these two languages may vary. Such differences are often ignored when considering the group as a whole.

Thus, we propose that using a recording system like LENA in different bilingual communities would be useful in examining how different bilingual infants around the world experience their language input differently. Indeed, some researchers have suggested that the socio-cultural aspects of a bilingual environment may be responsible for some of

the discrepancies in results across different phenomenon in bilingual research (Bayley & Schecter, 2003; Smithson, Paradis, & Nicoladis, 2014; Zentella, 1997). These socio-cultural factors may reveal themselves as differences in how bilingual caregivers talk to their child in different communities, which would be informative for describing bilingual acquisition in different contexts. Indeed, in our own research, we often cite how the bilingual community in Montréal is unique from that of other communities (for example, the Spanish-English community in the United States), in that French and English are both majority languages. However, it is time to take a step forward in describing how this difference might be reflected in the language input to a child growing up in Montréal versus a child growing up elsewhere. For example, our LENA findings show that overheard speech could be a rich source of dual-language input (often entailing more language mixing) for infants being raised in bilingual homes in Montréal. This may not describe infants in other bilingual contexts where language use is highly contextualized.

Certainly, the idea that infants' language experiences can vary widely is not a novel one. Many researchers acknowledge this variability by examining individual differences of language input in language learning. In our own study in Chapter 5, we provided evidence that, as a group, 8- and 10-month-old bilingual infants could segment words from both of their native languages. However, exploratory analyses suggest that it is those infants who heard more language mixing from their caregivers that were able to segment in both languages, while those who heard less language mixing demonstrated segmentation only in their native language. Thus, examining individual differences within bilingual groups would serve to increase our understanding of bilingual acquisition.

5.3. Future directions

As described above, our findings raise several issues in bilingualism, many of which can only be resolved through future research. With regards to research method assessment, our study examined different methodologies in investigating the language input in bilingual infants. First, one could examine how parental reports could be improved: What factors affect individuals' ability to track languages? How does the accuracy of these caregiver reports change at different age points? How can different interview formats improve individuals' ability to describe their own or their child's language environments? Second, one could examine how the LENA-generated measures could be used to examine the bilingual input: How do caregivers use their languages differently in different contexts? How do child vocalizations and turn-taking look like when they are exposed to multiple languages? Finally, one could explore different ways to efficiently analyze daylong recordings. To date, the LENA algorithms do not currently distinguish between different languages in the recordings. Developing a tool that is able to do so would help researchers efficiently analyze daylong recordings of bilingual infants. One could potentially use machine-learning approaches with our dataset to create and test such a tool, as our dataset contains a wide variety of bilingual speech (including code-mixing, accented speech, and both infant- and other-directed speech), and has been hand-coded for language.

With regards to examining the language input in bilingual infants, there are many future avenues to pursue. For example, our current findings suggest that it may be possible to efficiently gather quantitative aspects of a bilingual child's language input. In Chapter 2, we validated the LENA-generated estimates of overall word count to a bilingual

child, and in Chapter 3, we validated the parent-report measures of language exposure. This suggests that one could conduct language interviews with parents to estimate their child's language exposure ratio, and then multiply this ratio with the LENA-generated measures to obtain a volumetric estimate for each language. This measure could be very useful for examining different aspects of language acquisition – for example, comparing how absolute quantitative aspects of the language input affect outcomes between monolingual and bilingual infants, or examining the dose-response effects on speech processing in bilingual infants.

Our study focused on the variability of input quantity in bilingual infants, but one could also look at the quality of the input. Indeed, the quality of the input can differ in many dimensions. For example, some researchers have examined how infant-directed speech and conversational turns also affect language outcomes (Ramírez-Esparza et al., 2014; Romeo et al., 2018). Some recent studies have shown that these qualitative factors actually predict language outcomes better than quantitative measures of the language input (e.g., Dwyer, 2017; Hirsh-Pasek et al., 2015). It would be interesting to investigate how these two characteristics of the input are produced by bilingual caregivers, and how these characteristics affect infants' abilities in each language. In our own dataset, we are currently analyzing how some of these qualitative measures (e.g., turn-taking) might differ in different languages within the same person.

With regards to bilingual word segmentation, our study showed that 8- and 10-month-old bilingual infants have the processing capacity to segment bisyllabic words in their native languages, but the mechanism and nature of this ability has yet to be explored. Future questions include: What is the role of the amount of language exposure in bilingual

infants' ability to segment words? How would bilingual infants who are learning different pairs of languages perform in this task (especially those with structurally-different pairs of languages)? Does this bilingual advantage in word segmentation extend to unfamiliar languages? Answers to these questions will advance our understanding of speech processing and bilingual acquisition.

In addition to these questions, it would be interesting to examine how these variables change over time. We are currently following the 10-month-old bilingual infants tested in *Chapters 2, 3,* and *4* to examine how their language environment and language skills develop longitudinally. We have collected a second round of LENA recordings when they were 18-month-olds, in addition to examining their word learning skills in a laboratory task, and collecting parent-report measures of vocabulary development. These data will allow us to examine how characteristics of the language input change over time, including amount of language input from caregivers and siblings, proportion of exposure to each language, and amount of language mixing. Further, we can investigate how these measures of language input relate to their word learning skills and their vocabulary development. These data will contribute to the growing literature in giving us a clear picture on how children grow up to be bilingual.

6. References

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, *18*(1), 41–49. https://doi.org/10.1017/S0305000900013283

Andruski, J. E., Casielles, E., & Nathan, G. (2013). Is bilingual babbling language-specific? Some evidence from a case study of Spanish–English dual acquisition. *Bilingualism*, *17*(3), 660–672. https://doi.org/10.1017/s1366728913000655

Antovich, D. M., & Graf Estes, K. (2017). Learning across languages: bilingual experience supports dual language statistical word segmentation. *Developmental Science*, e12548. https://doi.org/10.1111/desc.12548

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 117–134.

Bayley, R., & Schecter, S. R. (2003). *Language Socialization in Bilingual and Multilingual Societies*. *Bilingual Education and Bilingualism*. Buffalo, NY: Multilingual Matters

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2018). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science,* e12715. https://doi.org/10.1111/desc.12715

Bergmann, C., & Cristia, A. (2018). Environmental Influences on Infants' Native Vowel Discrimination: The Case of Talker Number in Daily Life. *Infancy*, *23*(4), 484–501. https://doi.org/10.1111/infa.12232

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, *65*(4), 229–235. https://doi.org/10.1037/a0025406

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Blom, E. (2010). Effects of input on the early grammatical development of bilingual children. *International Journal of Bilingualism*, *14*(4), 422–446. https://doi.org/10.1177/1367006910370917

Bosch, L., Figueras, M., Teixidó, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, *4*, 106. https://doi.org/10.3389/fpsyg.2013.00106

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2–3), 217–243. https://doi.org/10.1177/00238309030460020801

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

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, *28*(3), 455–474. https://doi.org/10.1017/S0142716407070257

Byers-Heinlein, K. (2012). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism*, *16*(1), 32–48. https://doi.org/10.1017/s1366728912000120

Byers-Heinlein, K., & Lew-Williams, C. (2013). Bilingualism in the early years: What the

science says. *LEARNing Landscapes*, *7*(1), 95–112.

Byers-Heinlein, K., Schott, E., Gonzales-Barrero, A. M., Brouillard, M., Dubé, D., Jardak, A., … Tamayo, M. P. (submitted). Best practices for using parent report to measure bilingual infants' language environments.

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

Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2015). Reliability of the Language ENvironment Analysis system (LENA^TM) in European French. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-0634-8

Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2014). Adult Talk in the NICU With Preterm Infants and Developmental Outcomes. *Pediatrics*, *133*(3), e578–e584. https://doi.org/10.1542/peds.2013-0104

Charron, C., Fitzpatrick, E. M., McSweeney, E., Rabjohn, K., Somerville, R., & Steacie, P. (2016). Language ENvironment Analysis (LENA) with children with hearing loss: A clinical pilot. *Canadian Journal of Speech-Language Pathology & Audiology*, *40*(1), 92–104. https://doi.org/10.1016/j.jcomdis.2017.12.005

Core, C., & Hoff, E. (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

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: evidence from word segmentation. *Cognition*, *96*(3), 233–262.

https://doi.org/10.1016/j.cognition.2004.08.005

David, 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 Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, *28*(3), 411–424. https://doi.org/10.1017.S0142716407070221

De Houwer, A. (2009). Early bilingual acquisition. *Handbook of Bilingualism*, 30.

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

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Frienda, M. (2016). The Language Exposure Assessment Tool: Quantifying Language Exposure in Infants and Children. *American Journal of Speech-Language Pathology*, *25*, 1–15. https://doi.org/10.1044/2016

Donovan, W., Leavitt, L., Taylor, N., & Broder, J. (2007). Maternal sensory sensitivity, mother-infant 9-month interaction, infant attachment status: predictors of mother-toddler interaction at 24 months. *Infant Behavior and Development*, *30*(2), 336–352. https://doi.org/10.1016/j.infbeh.2006.10.002

Dwyer, A. R. (2017). Early language experience and later vocabulary among Australian infants from diverse socioeconomic backgrounds. (doctoral thesis). Western Sydney University, Penrith, Australia.

Dykstra, J. R., Sabatos-Devito, M. G., Irvin, D. W., Boyd, B. A., Hume, K. A., & Odom, S.

L. (2013). Using the Language Environment Analysis (LENA) system in preschool classrooms with children with autism spectrum disorders. *Autism*, *17*(5), 582–594. https://doi.org/10.1177/1362361312446206

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development*, *78*(5), 1510–1525. https://doi.org/10.1111/j.1467-8624.2007.01080.x

Fibla Reixachs, L. (2017) How could bilingual children segment their languages? A modeling approach (master's thesis). École normale supérieure, Paris, France.

Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2016.10.004

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. *Handbook of Language Development*, 324–342. https://doi.org/10.1002/9780470757833.ch16

Gilkerson, J., & Richards, J. A. (2009). The power of talk: Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development. *LENA Research Foundation Technical Report ITR-01-2*.

Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., … Topping, K. (2015). Evaluating language environment analysis system performance for Chinese: a pilot study in Shanghai. *Journal of Speech, Language and Hearing Research, 58*(2), 445–452. https://doi.org/10.1044/2015_JSLHR-L-14-0014

Gonzalez-Barrero, A. M., & Nadig, A. S. (2017). Can Bilingualism Mitigate Set-Shifting Difficulties in Children With Autism Spectrum Disorders? *Child Development*.

https://doi.org/10.1111/cdev.12979

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515–531. https://doi.org/10.1017/S0305000907008641

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Graf Estes, K., Gluck, S. C.-W., & Bastos, C. (2014). Flexibility in Statistical Word Segmentation: Finding Words in Foreign Speech. *Language Learning and Development*, 1–18. https://doi.org/10.1080/15475441.2014.926730

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, *15*(1), 77–94. https://doi.org/10.1017/S030500090001206X

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H Brookes Publishing.

Hart, B., & Risley, T. R. (1999). *The Social World of Children: Learning To Talk.* Baltimore, MD: Paul H Brookes Publishing.

Hart, B., & Risley, T. R. (2003). The Early Catastrophe. *Education Review*, *17*(1), 1–6. https://doi.org/ISSN-0148-432X

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., … Suma, K. (2015). The Contribution of Early Communication Quality to Low-Income Children's Language Success. *Psychological Science*, *26*(7), 1071–1083.

https://doi.org/10.1177/0956797615581493

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development*, *62*(4), 782–796. https://doi.org/10.2307/1131177

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

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

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Hohle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127. https://doi.org/10.1111/1467-7687.00261

Hohne, E. A., & Jusczyk, P. W. (1994). Two-month-old infants sensitivity to allophonic differences. *Perception & Psychophysics*, *56*(6), 613–623. https://doi.org/10.3758/BF03208355

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, *7*(3), 504–509. https://doi.org/10.3758/BF03214363

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236-248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Jusczyk, P. W., & Aslin, R. N. (1995). Infants detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1–23. https://doi.org/10.1006/cogp.1995.1010

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675–687. https://doi.org/10.1111/j.1467-8624.1993.tb02935.x

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, *61*(8), 1465–1476. https://doi.org/10.3758/BF03213111

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3), 159–207. https://doi.org/10.1006/cogp.1999.0716

Kay-Raining Bird, E., Lamond, E., & Holden, J. (2012). Survey of bilingualism in autism spectrum disorders. *International Journal of Language and Communication Disorders*, *47*(1), 52–64. https://doi.org/10.1111/j.1460-6984.2011.00071.x

Leimgruber, J. R. E. (2017). Global multilingualism, local bilingualism, official monolingualism: the linguistic landscape of Montreal's St. Catherine Street. *International Journal of Bilingual Education and Bilingualism*, 1–16. https://doi.org/10.1080/13670050.2017.1401974

LENA Research Foundation (2014). The LENA Research Foundation. Retrieved from

http://www.lenafoundation.org/.

Leopold, W. F. (1947). *Speech Development of a Bilingual Child, Vol. II: Sound-learning in the First Two Years.* Evanston, IL: Northwestern University Press.

Levitt, A. G., & Aydelott Utman, J. G. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, *19*(1), 19–49. https://doi.org/10.1017/S0305000900013611

Li, L., Vikani, A. R., Harris, G. C., & Lin, F. R. (2014). Feasibility study to quantify the auditory and social environment of older adults using a digital language processor. *Otology and Neurotology*, *35*(8), 1301–1305. https://doi.org/10.1097/MAO.0000000000000489

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431.

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, *6*(5), 314–317. https://doi.org/10.1111/j.1467-9280.1995.tb00517.x

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2016). Caregiver talk to young Spanish-English bilinguals: comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*. https://doi.org/10.1111/desc.12425

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–494. https://doi.org/10.1006/cogp.1999.0721

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information

can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111. https://doi.org/10.1016/S0010-0277(01)00157-3

McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, *95*(2), B15-26. https://doi.org/10.1016/j.cognition.2004.07.005

Mueller Gathercole, V. C., Thomas, E. M., Roberts, E. J., Hughes, C. O., & Hughes, E. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In *Issues in the Assessment of Bilinguals*.

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*(1), 95–120. https://doi.org/10.1017/S0305000997003358

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), 756-766. https://doi.org/10.1037/0096-1523.24.3.756

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283–299. https://doi.org/10.1016/j.jml.2005.10.004

Nelson, D. G. K., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*(1), 111–116. https://doi.org/10.1016/0163-6383(95)90012-8

Newman, R., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006).

Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology*, *42*(4), 643–655. https://doi.org/10.1037/0012-1649.42.4.643

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2015). Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, 1–16. https://doi.org/10.1017/S0305000915000446

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., … Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, *107*(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107

Orena, A.J., Srouji, J. S., Byers-Heinlein, K., & Polka, L. (2018). Evaluating the LENA recording system for investigating speech input in a French-English bilingual context , Acoustical Society of America, May 2018 [poster session], Minneapolis, MN.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn Psychol*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J. A., Ma, L., & Topping, K. (2016). Effects of feedback on parent–child language with infants and toddlers in Korea. *First Language*, *36*(6), 549–569. https://doi.org/10.1177/0142723716649273

Paradis, J. (2007). Bilingual children with specific language impairment: Theoretical and applied issues. *Applied Psycholinguistics,* 28, 551–564.

https://doi.org/10.1017/S0142716407070300

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

Patterson, J. L. (1998). Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *American Journal of Speech-Language Pathology*, *7*(4), 46–56. https://doi.org/10.1044/1058-0360.0704.46

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41–58. https://doi.org/10.1017/S0142716400009863

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

Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2017). Segmenting words from fluent speech during infancy – challenges and opportunities in a bilingual context. *Developmental Science*, *20*(1). https://doi.org/10.1111/desc.12419

Polka, L., & Sundara, M. (2012). Word Segmentation in Monolingual Infants Acquiring Canadian English and Canadian French: Native Language, Cross-Dialect, and Cross-Language Comparisons. *Infancy*, *17*(2), 198–232. https://doi.org/10.1111/j.1532-7078.2011.00075.x

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

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2016). The Impact of Early Social Interactions on Later Language Development in Spanish-English Bilingual Infants. *Child Development*. https://doi.org/10.1111/cdev.12648

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

Rowe, M. L., Leech, K. A., & Cabrera, N. (2016). Going Beyond Input Quantity: Wh-Questions Matter for Toddlers' Language and Cognitive Development. *Cognitive Science, 41*(S1), 162-179. https://doi.org/10.1111/cogs.12349

Rvachew, S., Rees, K., Carolan, E., & Nadig, A. (2017). Improving emergent literacy with school-based shared reading: Paper versus ebooks. *International Journal of Child-Computer Interaction*, *12*, 24–29. https://doi.org/10.1016/j.ijcci.2017.01.002

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science, 274*(5294), 1926-1928. https://doi.org/10.1126/science.274.5294.1926

Seidl, A., Cristià, A., Bernard, A., & Onishi, K. H. (2009). Allophonic and Phonemic Contrasts in Infants' Learning of Sound Patterns. *Language Learning and Development*, *5*(3), 191–202. https://doi.org/10.1080/15475440902754326

Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*(2), 128–142.

https://doi.org/10.1016/j.cognition.2012.05.008

Singh, L., Reznick, S. J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, *15*(4), 482–495. https://doi.org/10.1111/j.1467-7687.2012.01141.x

Slonims, V., & Mcconachie, H. (2006). Analysis of Mother–Infant Interaction in Infants With Down Syndrome and Typically Developing Infants. *American Association on Mental Retardation*, *273*(4), 273–289. https://doi.org/10.1352/0895-8017(2006)111[273:AOMIII]2.0.CO;2

Smithson, L., Paradis, J., & Nicoladis, E. (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Bilingualism*, *17*(4), 810–821. https://doi.org/10.1017/S1366728913000813

Statistics Canada (2016). Focus on Geography Series, 2016 Census. *Statistics Canada Catalogue no. 98-404-X2016001*. Ottawa, Ontario.

Stromswold, K. (2001). The Heritability of Language: A Review and Metaanalysis of Twin, Adoption, and Linkage Studies. *Language*, *77*(4), 647–723. https://doi.org/10.1353/lan.2001.0247

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*(2), 454-464. https://doi.org/10.1037/0012-1649.43.2.454

Tamis-LeMonda, C. S., & Bornstein, M. H. (2002). Maternal responsiveness and early language acquisition. *Advances in Child Development and Behavior*, *29*, 90–129. https://doi.org/10.1016/S0065-2407(02)80052-0

Tamis-LeMonda, C. S., Bornstein, M. H., Baumwell, L., & Melstein Damast, A. (1996).

Responsive parenting in the second year: Specific influences on children's language and play. *Early Development and Parenting*, *5*(4), 173–183. https://doi.org/10.1002/(SICI)1099-0917(199612)5:4<173::AID-EDP131>3.0.CO;2-V

Tamis-Lemonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2018). Routine Language: Speech Directed to Infants During Home Activities. *Child Development*. https://doi.org/10.1111/cdev.13089

Tamis-Lemonda, C. S., Kuchirko, Y., & Song, L. (2014). Why Is Infant Language Learning Facilitated by Parental Responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Thiemann-Bourque, K. S., Warren, S. F., Brady, N., Gilkerson, J., & Richards, J. A. (2014). Vocal interaction between children with down syndrome and their parents. *American Journal of Speech-Language Pathology*, *23*(3), 474–485. https://doi.org/10.1044/2014_AJSLP-12-0010

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

Tsui, A., Erickson, L. C., Thiessen, E., & Fennell, C. (2017). Statistical Learning from Accented Speech: A Bilingual Advantage. *Proceedeings of the 41st Annual Boston University Conference on Language Development, 679 - 690.* Sommerville, MA: Cascadilla Press.

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In *Assessing Multilingual Children: Disentangling Bilingualism from Language*

Impairment.

Van de Weijer, J. (1999). Language input for word discovery (doctoral thesis. Radboud University Nijmegen Nijmegen.

van Heugten, M., & Johnson, E. K. (2017). Input matters: Multi-accent language exposure affects word form recognition in infancy. *The Journal of the Acoustical Society of America*, *142*(2), EL196-EL200. https://doi.org/10.1121/1.4997604

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., & Plummer, F. (2012). Parent-infant interaction in infant siblings at risk of autism. *Research in Developmental Disabilities*, *33*(3), 924–932. https://doi.org/10.1016/j.ridd.2011.12.011

Weisleder, A., & Fernald, A. (2013). Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: a challenge for statistical learning? *Language Learning and Development*, *5*(1), 30–49. https://doi.org/10.1080/15475440802340101

Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics*, *83*, 101.

Wermke, K., Ruan, Y., Feng, Y., Dobnig, D., Stephan, S., Wermke, P., … Shu, H. (2017). Fundamental Frequency Variation in Crying of Mandarin and German Neonates. *Journal of Voice*, *31*(2), 255.e25-255.e30. https://doi.org/10.1016/j.jvoice.2016.06.009

Wood, C., Diehm, E. A., & Callender, M. F. (2016). An Investigation of Language

Environment Analysis Measures for Spanish-English Bilingual Preschoolers From Migrant Low-Socioeconomic-Status Backgrounds. *Language, Speech, and Hearing Services in Schools*, *47*(2), 123–134. https://doi.org/10.1044/2015_LSHSS-14-0115

Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York*. Malden, MA : Blackwell Publishers.

Zhao, T., Moon, C., Lagercrantz, H., & Kuhl, P. (2011). Prenatal Motherese? Newborn Speech Perception May Be Enhanced by Having a Young Sibling. *Psi Chi Journal of Undergraduate Research*, *16*(2).