Speech Disfluencies: Their Role in Comprehension and Word Learning

Elizabeth Morin-Lessard

A Thesis
In the Department
of
Psychology

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

August 2018

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Elizabeth Morin-Lessard

Entitled: Speech Disfluencies: Their Role in Comprehension and Word Learning

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

Doctor of Philosophy (Psychology)

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

Signed by the final examining committee:

Dr. Elyse Amend	Chair
Dr. Katherine S. White	External Examiner
Dr. Pavel Trofimovich	External to Program
Dr. Diane Poulin-Dubois	Examiner
Dr. Norman Segalowitz	Examiner
Dr. Krista Byers-Heinlein	Thesis Supervisor

Approved By:

Dr. Andrew Chapman, Graduate Program Director

October 3, 2018

Dr. André Roy, Dean of the Faculty of Arts and Sciences

Abstract

Speech Disfluencies: Their Role in Comprehension and Word Learning

Elizabeth Morin-Lessard, Ph.D.
Concordia University, 2018

One of the most extraordinary aspects of human development is how children acquire their language(s) by listening to spontaneous speech. Perhaps more remarkably, they do so even though speech is often highly disfluent. To better understand how language acquisition unfolds, this dissertation explored the effects of speech disfluencies on real-time word comprehension and on word learning for listeners from different language backgrounds and levels of expertise: monolingual children, bilingual children, and bilingual adults.

Manuscript 1 reports two word comprehension studies, looking at the ability of children and adults to use disfluencies to predict whether a speaker will name a novel or a familiar object. This ability was investigated by presenting sentences with disfluencies in listeners' native and non-native language(s). Study 1 tested 32-month-old monolingual and bilingual children, and Study 2 tested bilingual adults. Results from Studies 1 and 2 indicate that listeners looked more at the novel than the familiar object upon hearing a disfluency, irrespective of participants' language experience, and whether the disfluency was in participants' native language(s). Importantly, the results suggest that listeners might attend to a speaker's uncertainty more than the particular realization of the disfluency.

Manuscript 2 investigates the impact of speech disfluencies on novel word learning in monolingual and bilingual 32-month-old children. We considered two contrasting possibilities: (1) Disfluencies will facilitate novel word learning, since listeners direct looks to novel objects upon hearing a disfluency, versus (2) disfluencies will hinder novel word learning, since they signal a speaker's uncertainty about an object's label. The results indicate that disfluencies may hinder novel word learning: Children did not learn the novel words following disfluencies, nor the novel words following fluent speech. Though somewhat inconclusive as children did not learn words in either case, these results suggest that children's word learning may be hindered when a speaker is disfluent.

Together, the results from the two manuscripts in this dissertation suggest that speech disfluencies are a double-edge sword: they can be helpful for making predictions during real-time comprehension, but could hinder word learning. These findings have important implications for understanding the role of speech disfluencies in language acquisition.

Acknowledgements

First and foremost, I would like to express my infinite gratitude to my supervisor Dr. Krista Byers-Heinlein for the time and energy she has invested in me throughout my doctoral studies. Her enthusiasm and rigor for research have greatly enriched my Ph.D. experience, and have motivated me to become a better scientist and writer. Throughout the dissertation-writing period, she provided encouragement and always helped me see the first steps of an at times daunting process. I am also thankful for her invaluable professional development advice, and for being a role model by embodying feminist values both in her mentorship style and in the workplace. Krista, it has truly been an honour for me to be your very first graduate student.

I would also like to thank the members of the Bilingualism Interest Group (BIG) at Concordia University, who made several useful suggestions for my work and asked questions to challenge me as a researcher. Particular thanks to my committee members Dr. Diane Poulin-Dubois and Dr. Norman Segalowitz for their interest in my work, and for their insightful comments during the planning and writing stages of this dissertation.

I also want to extend my thanks to developmental researchers Dr. Katherine White and Dr. Celeste Kidd for insightful design discussions on Manuscript 1, and for sharing their audio stimuli with me. Their input has highly valuable for replicating their original speech disfluencies work.

Completing my research would have been substantially more difficult without the many research discussions and valuable encouragements of past and current members of the Concordia Infant Research Lab: Sonia Giulione**,** Michela Martiniello, Madeleine Borgeat**,** Sadaf Pour Iliaei, Maria Tamayo, Meghan Mastroberardino, Nicholas Salama-Siroishka, and Alexandra Laoun-Rubenstein. A first big thank you goes to Kay Goly for her assistance with stimuli preparation and data collection for Manuscript 1. Particular thanks to my punny friend and lab manager extraordinaire Melanie Brouillard for brightening my days, to Amel Jardak for never failing to make me laugh (especially with discussions of feudalism), and to Chelsea da Estrela for always being up for a quick coffee or an adventure. Thank you also to Esther Schott and Daphnée Dubé for their friendship and inspiring work ethic. Although I have only had the chance to know Dr. Ana Maria Gonzalez Barrero for a short time, I am very appreciative of her support and perspective in the final stages of my Ph.D.. It's been a real pleasure to work and constantly *code-switch* with you all.

Big thanks go out to my fellow graduate students who became friends over the years. Thank you to my favourite Ottawa U. syntactician and roommate Nova Starr for being such a great friend and cheerleader, and to my Concordia U. office mate and cohort friend Dr. Mélanie Béland for the recent motivation to follow her footsteps in defending my dissertation. I could not forget to thank my fellow Ph.D. students at McGill University Dr. Adriel John 'AJ' Orena, Dr. Amélie Bernard, and Dr. Sarah Colby. Our regular R and writing meetings at Kaféin and Anticafé – particularly in the final writing stages of this dissertation – have kept me motivated and somewhat social.

I would not be where I am today without the guidance and support of my former supervisors. Many thanks to Bob Dylan's #1 Fan, Dr. Stuart J. McKelvie, for being the first to believe in my potential as a researcher, and for being an incredible mentor since my graduation from Bishop's University. I am also very thankful to my M.A. supervisor Dr. Tania S. Zamuner for her guidance, and for giving me amazing work and research opportunities, both while under her supervision and in the years beyond my time at the University of Ottawa.

My research would not have been possible without the financial help of external and internal funding sources, nor without participants. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (402470-2011) to Krista Byers-Heinlein, and by a fellowship from the Fonds de Recherche du Québec – Société et Culture to Elizabeth Morin-Lessard. Dissemination and presentation of this work at conferences was also made possible thanks to financial support from the Department of Psychology, the Faculty of Arts and Science, and the School of Graduate Studies at Concordia University. Many thanks to all the adults and the families of children who took part in my studies, and provided rich and interesting data for me to discuss in this dissertation.

Finally, I want to give the most heartfelt thanks to my family and friends for their love, patience, and encouragements throughout this challenging process of earning my doctorate degree. Maman, Marie-Eve, Louis-Philippe, Marie-Philipp, *et* Daniela, *je suis choyée de vous avoir dans ma vie et de pouvoir compter sur votre soutien indéfectible, tant pour mes décisions professionnelles que personnelles.*

To all of you who have one way or another contributed in the completion of this dissertation: *Merci!*

Dedications

To my dog Zelda, and coffee. Though you have often deprived me of some precious sleep, you have kept me sane and given me instant comfort on more occasions than I can count.

Contribution of Authors

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

Manuscript 1 (CHAPTER 2; comprising 2 studies)

Morin-Lessard, E., & Byers-Heinlein, K. (*under review*). Uh and euh signal novelty for monolinguals and bilinguals: Evidence from children and adults. Manuscript submitted for publication.

Manuscript 2 (CHAPTER 3; comprising 1 study)

Morin-Lessard, E., & Byers-Heinlein, K. (*in preparation*). Word learning following speech disfluencies in monolingual and bilingual children.

Relative Contributions

For both manuscripts, I developed the original research idea, designed the studies, created the stimuli, tested participants, cleaned the data for analysis, performed the analyses, interpreted the results, and drafted and revised the manuscripts. Honours student Kay Goly (Manuscript 1) and other research assistants at the Concordia Infant Research lab (Manuscripts 1 and 2) helped with participant testing. Dr. Krista Byers-Heinlein provided feedback on study design and stimuli, statistical analyses, interpretation of the results, and manuscript drafts.

Table of Contents

List of Figures

CHAPTER 2: Manuscript 1

CHAPTER 3: Manuscript 2

List of Tables

CHAPTER 2: Manuscript 1

CHAPTER 3: Manuscript 2

1 CHAPTER 1: General Introduction

Consider the following sentence: "What is the date today?" If you were to give this perfectly enunciated voice command to your smart phone, it would likely give you today's date. Now, try it using a more natural version of the same sentence, this time including the common speech disfluency *uh*, for example: "What is the *uh*… date today?" If your smart phone is equipped with the most recent voice recognition software, you will note that 1) *uh* has been edited out, and 2) your phone will give you today's date. Through a complex sequence of analyses and machine learning algorithms designed by scientists and programmers (Duchateau, Laureys, Demuynck, & Wambacq, 2003; Siri Team, 2017), voice recognition software seems to seamlessly handle complex natural speech by recognizing *uh,* before discarding it to process the remainder of the sentence.

Now consider being a young English-learning child with no formal education, and certainly no programming experience. As a child, your experience with language mainly comes from your caregivers' natural speech, which is highly disfluent (Shriberg, 2001; Fox Tree, 1995). If you heard the sentence "Where is the *uh*… narwhal?", one option would be to discard the speech disfluency, just like voice recognition software. Perhaps more advantageously, a second option would be to use the speech disfluency *uh* in order to correctly identify the unfamiliar whale among other familiar objects. Research with monolingual English children indicates that instead of discarding speech disfluencies, children use them to their advantage during comprehension. For example, when English-learning children hear "Look! Look at the *uh*…", they predict that the speaker will label a novel object instead of a familiar one, before even hearing a specific object label (Kidd, White, & Aslin, 2011). One reason why children are thought to look at a novel object upon hearing speech disfluencies is because speech disfluencies tend to be produced before words that are new to the conversation (Arnold, Losongco, Wasow, & Ginstrom, 2000). Monolingual English children appear to notice that disfluencies often precede words that are new to the conversation (which may also be novel for children), making speech disfluencies a useful cue for these children to make predictions about upcoming words during listening. Reasons why disfluencies can serve as a useful cue during listening will be discussed further in the Speech Disfluencies in Communication section of this general introduction.

English monolingual children can use speech disfluencies to make predictions about a likely upcoming word. What about children of different language backgrounds, such as bilingual children? Indeed, both monolingual and bilingual children hear speech disfluencies in the speech of their caregivers: in one language for monolingual children, and in two languages for bilingual children. For example, monolingual English-learning children hear the English speech disfluency *uh*. In comparison, bilingual children exposed to English and French likely hear both the English speech disfluency *uh* in English sentences, and the French speech disfluency *euh* in French sentences. Because their caregivers are often also bilingual, these children may also hear the French speech disfluency *euh* in English sentences, and vice-versa. Although monolinguals and bilinguals both face the challenge of navigating speech that is disfluent, their experience with speech disfluencies is fundamentally different. Yet, it is not clear whether speech disfluencies from different languages will be equally useful to young listeners to predict upcoming speech.

The research reported in this dissertation harnesses bilingualism as a lens for understanding speech disfluencies, to shed light on how they are processed and used by young children and adults who hear them in spontaneous speech. In this dissertation, I sought to expand the current understanding of disfluencies by answering the following question: How do listeners from monolingual and bilingual backgrounds comprehend and learn words in disfluent speech? This introductory chapter reviews the central background literature relevant to this question. In section 1.1, I first provide a general overview of early language acquisition and discuss two important mechanisms involved in language acquisition: word learning and word comprehension. In section 1.2, I provide a definition of speech disfluencies, and discuss their role in communication. In section 1.2.1, I review the extant literature on the role of speech disfluencies in monolingual children's language comprehension, and in section 1.2.2 I discuss the potential role of speech disfluencies during word learning. In section 1.3, I discuss bilingual language acquisition, and argue why studying bilinguals can provide a unique window into the relationship between speech disfluencies and language acquisition. Finally, in section 1.4, I outline the main objectives of this dissertation, and how the two manuscripts following this introductory section address those objectives.

1.1 Language Acquisition

Language is an exceptionally complex and unique ability that allows humans to understand and express an infinite number of ideas. Nevertheless, every typically-developing child goes through the process of acquiring one or more native languages, and eventually becomes a successful language user. While the general process of language acquisition is universal, children's particular experience acquiring language may vary: Some children are exposed to and learn only one language and become monolingual, while others are exposed to and learn two languages and become bilingual. For both monolingual and bilingual children, language acquisition begins in children's earliest experiences hearing their mother's voice in the womb (Byers-Heinlein, Burns, & Werker, 2010), and continues as they begin to understand, learn, and produce words in their language(s). Children can understand common words (e.g., banana) from as early as 6 months (Bergelson & Swingley, 2012), and begin to produce their first words around their first birthday (Fenson et al., 2007). Thus, in a matter of just a few months, children begin to comprehend and learn words from the speech stream, and acquire one, two, or more languages with remarkable success by listening to spontaneous speech.

In this dissertation, I focus on two milestones of language acquisition: word learning and real-time word comprehension. Word learning refers to the process by which learners encode relationships between auditory word forms and their referents into memory, for example linking the word 'ball' with the representation of a solid spherical object. Real-time word comprehension refers to the ability to recognize already learned, familiar words in the context of running speech, for example by recognizing the word *ball* in the sentence "Look at the red ball!". Both of these abilities begin to develop in the first year of life, in conjunction with several other prerequisite abilities. For instance, very young infants show evidence of speech perception abilities: they prefer their mother's voice over the voice of other females (Mehler et al., 1978; DeCasper & Fifer, 1980), and prefer the familiar nature of their native language(s) over unfamiliar languages (e.g., Byers-Heinlein et al., 2010; Moon, Cooper, & Fifer, 1993; Nazzi, Jusczyk, & Johnson, 2000). This early sensitivity to familiar voices and languages can be helpful for infants to tune into the sounds of their native language(s), and to collect important information about the speech sounds that make up words in such language(s). Although infants show an early sensitivity to familiar speech, infants are born able to discriminate sound contrasts from any language. However, with increased exposure to their native language in the first year of life, children become better at discriminating between native sound contrasts (Kuhl et al., 2006) at the expense of non-native sound contrasts (Werker & Tees, 1984; and see review by Werker, Yeung, & Yoshida, 2012).

More relevant to this dissertation is how infants process speech at the word level. Infants must locate individual word units from the speech stream (e.g., determiners, nouns, verbs) before they can comprehend or learn them, which can be challenging since, unlike in written language, there are no spaces or pauses between words in spontaneous speech. To determine where one word ends and the next begins, children must be able to notice that certain syllables tend to co-occur and assume that these are likely words (Saffran, Aslin, & Newport, 1996), or to notice that some syllables are stressed and others are not (e.g., syllables are stressed at the beginning of words in English; Jusczyk, Houston, & Newsome, 1999). Thus, within a year of being born, children develop several abilities, which shape how they process sounds and words in their native and non-native languages. These abilities contribute to facilitating children's learning and comprehension of words in the language(s) relevant to them. The next two sections are dedicated to these two important aspects of language acquisition: word learning and word comprehension.

1.1.1 Word learning

When a child locates a word in the speech stream, that word may be either novel or familiar to them. In this section, I will discuss word learning, which takes place when children encounter novel words; in the next section, I will discuss word comprehension, which takes place when children encounter familiar words. Novel words represent a word learning opportunity for children. Minimally, word learning requires children to determine what a word refers to, and store this word with its referent in memory. In the context of word learning, a referent is the object to which a word *refers* (as opposed to merely be associated with; Waxman & Gelman, 2009). For example, children hearing "Look at the narwhal!" would have to process and encode the sounds of the word 'narwhal', select the likely referent of the word, and store the sound-referent pair together in memory. With exposure to different narwhals, children eventually store a new sound-referent pair along with other representations of similar entities and concepts (e.g., other similar-looking animals which are also called 'narwhal'). Word learning is an important skill for young children who use this mechanism to build their vocabularies, and eventually become proficient speakers. In the following paragraphs, I will first discuss how children can select the likely referent of a novel word, and I will then discuss retention, which involves the storing of the novel word-referent in memory.

Determining the likely referent of a novel word is not a trivial task. Here, I will focus on noun learning. Children see many objects around them, and must first determine which particular object among the ones they see the word ‘narwhal’ refers to. The identification of the appropriate referent for a particular novel word that children hear is crucial for word learning. Indeed, associating a novel word with the wrong object or failing to associate a novel word with a novel object at all might result in listeners not knowing what a speaker is talking about. How do children successfully identify the correct referent for a particular novel word they hear? Two ways in which they can do so is by relying on social cues from speakers (Brooks & Meltzoff, 2002; Colonnesi, Stams, Koster, & Noom, 2010; Emery, 2000; Yow & Markman, 2016) and on contextual cues offered by objects in the environment (e.g., familiar vs. novel objects; Byers-Heinlein & Werker, 2009; 2013; Davidson & Tell, 2005; Houston-Price, Caloghiris, & Raviglione, 2010).

Social cues from speakers can be a rich source of information to help children determine the likely referent of a novel word. For example, speakers often gaze or point towards the object that they are referring to, and young infants become sensitive to these ostensive cues in their second year of life (Brooks & Meltzoff, 2002; Colonnesi, Stams, Koster, & Noom, 2010; Emery, 2000; Hollich, Hirsh-Pasek, & Golinkoff, 2000; Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Yow & Markman, 2016). Cues such as pointing and gazing towards objects are very effective in helping children know which word speakers refer to, as they provide direct reference information to a particular object in space. Using ostensive cues, children can know what a speaker is referring to from very early on, and hone their skills as they get older. Many researchers take such abilities as evidence that young listeners are trying to discern what a speaker is attempting to communicate, which can facilitate speech comprehension since they will know what the speaker is talking about (see review by Tomasello, 2000).

If clear social cues are not available during word learning, children still have several other word learning heuristics available to help them guess a new word’s potential referent. One such heuristic relevant to this dissertation, called mutual exclusivity, is the assumption that objects should only have one label (Markman & Wachtel, 1988). Monolingual children are particularly good at using this assumption, and show mutual exclusivity from as young as 17 months (Halberda, 2003). For example, when presented with a novel object (e.g., a cherry pitter) and a familiar object (e.g., a spoon), and asked to find an object with an unfamiliar novel name (e.g., “Show me the x!”, where x is a nonsense word), monolingual children tend to choose the novel object as the most likely referent (Markman & Wachtel, 1988). Thus, monolingual children can use their knowledge of a familiar word to infer the reference of a novel word. This is trickier for bilingual children, who usually have two labels for the same object: one in each of their languages. As a result, bilingual children are typically less consistent in their use of mutual exclusivity than monolingual children (Byers-Heinlein & Werker, 2009; 2013; Davidson, Jergovic, Imami, & Theodos, 1997; Davidson & Tell, 2005; Houston-Price, Caloghiris, & Raviglione, 2010). Children's ability to infer that the novel object, rather than the familiar one, is the most likely referent for a novel name is an important assumption for word learning. The assumption of mutual exclusivity is thus a useful strategy that children can use to identify a referent and learn words, though this strategy may not necessarily be used in the same way or to the same extent by monolinguals and bilinguals.

Once children have identified the referent of a novel word – whether via social information or their existing knowledge of words and objects – the hard work of word learning is not complete. Establishing reference of a word to an object may not be sufficient for word learning to take place. In fact, there is mounting evidence that sometimes children are able to determine a word's likely referent, but do not store the novel word with its referent in memory. Research indeed suggests that there may be a dissociation between using mutual exclusivity to select the referent of a novel label, and actually linking the referent and the label. That is, sometimes children use mutual exclusivity to gaze at a novel referent upon hearing a novel label, but do not appear to have retained that mapping when tested later (Bion, Borovsky, & Fernald, 2013; Horst & Samuelson, 2008; Samuelson & McMurray, 2017). Thus, retaining a word-object combination in memory and storing it along with other similar representations and concepts is much more difficult than the single step of selecting the appropriate object for a word.

Some factors facilitate retention of novel words with their referents in memory. For example, word retention is enhanced by prior exposure to a word-object pairing (Swingley, 2007). Indeed, repeated presentation of a word in the presence of its referent has been found to strengthen children's memory for a novel word (see Gathercole, 2006 for a review). Moreover, salient and interesting objects (e.g., bright, colorful, moving, and novel objects) are more likely to capture children's attention than boring objects, and children are more likely to retain a label for such objects (particularly in children's first two years of life; e.g., Hollich et al., 2000; Houston-Price, Plunkett, & Duffy, 2006; Pruden et al., 2006; Samuelson & Smith, 1998; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). The word form itself also matters for word learning, as some linguistic forms may be considered by children as better likely candidates referring to an object than other sounds. For example, children aged about 12 months are more likely to associate well-formed words with objects than non-word communicative sounds (e.g., *mmm* and *shhh*; MacKenzie, Graham, & Curtin, 2011), and words with sounds that are plausible in their native language with objects compared to words with sounds that would not be plausible in their native language (MacKenzie, Curtin, & Graham, 2012).

In sum, two dissociable individual processes are minimally involved in word learning: referent selection and word-referent retention. Both mechanisms are involved when children encounter novel words and engage in word learning. If any steps prior to the retention of the word-object association are skipped or reference to the correct object is fragile, a novel word may not be properly learned. Fortunately, children can make use of heuristics offered by speakers and objects to facilitate the identification of the correct referent for a novel word, and can rely on properties of language and objects to facilitate the retention of the novel word and its referent in memory.

1.1.2 Word comprehension

Once novel words have been learned and stored in memory, they become familiar words and word comprehension can occur. To comprehend a familiar word in running speech (i.e., in real time), children need to hear the speech unit, recognize it as a word that they know, map it to its existing representation as a familiar word in memory, access this word's existing object representation, and retrieve the word-object representation from memory. This process cannot occur when hearing novel words, as there is no familiar word to recognize: Children cannot comprehend novel words by accessing existing word-object representations in memory, as they do not yet have an existing representation. In the paragraphs below, I will discuss the development of real-time speech comprehension abilities, and how comprehension is facilitated by predictive processes.

In everyday conversation, speech occurs at a rapid rate, and listeners must comprehend this speech as it occurs in real-time. Efficient real-time speech processing contributes to better comprehension abilities. For example, adults begin to process a word as soon as they hear the first sounds rather than waiting to hear the entire word (Marslen-Wilson & Zwitserlood, 1989), which yields faster access to its meaning. Fast access to the meaning of words is an important skill for young listeners that enables them to follow and comprehend speech as in unfolds.

Infants make important gains in word processing and comprehension within their first years of life, and become increasingly fast and accurate at recognizing familiar words between 15 and 24 months (Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998). Moreover, infants show links between efficient real-time comprehension and vocabulary size (Legacy, Zesiger, Friend, & Poulin-Dubois, 2016) and growth (Fernald & Marchman, 2012). For example, typically-developing children who were categorized as late-talkers but were more efficient in word processing at 18 months showed faster vocabulary growth compared to late-talkers who were less efficient in word processing at the same age (Fernald & Marchman, 2012). Fast processing is also related to later language development (e.g., lexical and grammatical skills, working memory) beyond children's second year of life (Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008). The ability to quickly identify and process words therefore provides an advantage for young learners, as it facilitates real-time comprehension, which in turn supports word learning.

Another factor that facilitates word comprehension is the ability to use one's language experience and the cues available in language itself. For example, being able to predict an upcoming word based on the previous sentence structure or based on verbs (e.g., the verb "eat" is more likely to precede a food item than a non-food item; Friedrich & Friederici, 2005; Snedeker & Trueswell, 2004) facilitates word comprehension. Listeners' existing knowledge about word predictability and a likely upcoming word can facilitate the processing of new sentences, thereby reinforcing previously-learned representations of likely upcoming words for a particular sentence structure (Conway, Bauernschmidt, Huang, & Pisoni, 2010).

While the focus of this section has been on familiar word comprehension, it is important to note that processes involved in word comprehension can in turn influence word learning processes. Typically, listeners encounter a novel word in the context of a sentence that also contains familiar words. Using the example of the narwhal again, a child might hear his or her caregiver say "Look at the dog, the pig, and the narwhal!". In this case, assuming that a child already has a word and referent for dog and pig and successfully retrieves these words from memory, this child might assume that narwhal is a novel word that has never been heard before and be able to identify the noun as novel. Similarly, because this child knows what dogs and pigs typically look like but is seeing a narwhal for the first time, he or she might be able to identify the novel object referent in a display. In sum, the learning of novel word forms from the speech input is heavily influenced by listeners' pre-existing linguistic knowledge of familiar words and sentences, and this knowledge of familiar forms (e.g., dog, pig) and of novel ones (e.g., narwhal) gets strengthened with more exposure to incoming information from the speech stream.

1.2 Speech Disfluencies in Communication

Most of the studies that have investigated word learning and word comprehension have presented children with seamless fluent speech. For example, word comprehension and word learning studies typically present children with fluent sentences such as "Look at the X!" or "Can you see the X?" (e.g., Byers-Heinlein & Werker, 2009; Fernald, Swingley, & Pinto, 2001). However, natural speech is rarely perfectly fluent, and children must comprehend speech and learn new words in the context of disfluent speech. In the next section, I will turn to a discussion of where speech disfluencies are produced in communication, and discuss their characteristics. Following this, I present evidence that speech disfluencies can help with children's real-time speech comprehension, and discuss how they may also be involved in children's word learning.

Speech disfluencies, for example filled pauses produced by speakers such as *uh* and *um,* serve a role in communication. They are produced in various forms: As editing terms (e.g., "I don't know", "I guess"), silent pauses, and filled pauses (e.g., *uh*, *um*) (Smith & Clark, 1993). The term *disfluency* is derived from a combination of the words "impaired" and "flow". In this dissertation, I will refer to disfluent speakers as generally proficient speakers from non-clinical populations. This is because, unlike proficient speakers, speakers from clinical populations, for example individuals with cognitive impairments (e.g., neurological damage), may experience an impaired flow of speech and produce more speech disfluencies compared to healthy individuals (Yairi, Gintautas, & Avent, 1981; Lundgren, Helm-Estabrooks, & Klein, 2010). For most individuals who have not been diagnosed with a language impairment, however, the production of speech disfluencies is not pathological. In fact, speech disfluencies are thought to be universal. Disfluencies have indeed been observed in the speech of highly proficient speakers (Cutler, 1981), including speakers of high status (Hopper, 2016).

The evidence that speech disfluencies are produced even in speakers of high status – who likely deliver well-rehearsed addresses – suggests that they are the product of a linguistic difficulty that was not planned. When produced, disfluencies allow speakers to repair speech mistakes made along the way, and serve pragmatic functions such as allowing speakers to keep their turn in conversation (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001), and to intentionally (Clark & Wilkes-Gibbs, 1990) or unintentionally (Finlayson & Corley, 2012) communicate a linguistic difficulty to listeners. Depending on when, where, and with whom they are produced, speech disfluencies are manifested in many forms in natural speech (e.g., editing terms, silent pauses, and filled pauses; Smith & Clark, 1993). The focus of this dissertation is the filled pause, specifically *uh*, one of the most common speech disfluencies in English (Shriberg, 2001), and its French equivalent *euh*. In this dissertation, such filled pauses will be referred to using the more general terms 'speech disfluency' or simply 'disfluency'.

Despite the apparent randomness of speech disfluencies in spontaneous speech, they are systematic both in their occurrence and their form. Analyses of speech disfluencies have revealed regularities in when they are produced. Speakers may experience difficulties and produce disfluencies when planning a complex sentence, evidenced by the fact that disfluencies tend to occur more before longer sentences (Shriberg, 1996). Speakers also produce disfluencies when they are thinking about the most appropriate word (Vitkovitch & Tyrrell, 1995) or message (Fox Tree & Clark, 1997) to use in conversation. Moreover, speakers produce disfluencies when suppressing a competing word or response (Engelhardt, Nigg, & Ferreira, 2013), or to signal to listeners that they are experiencing a difficulty, and that help might be needed (Clark & Wilkes-Gibbs, 1990). The type of disfluency used may also vary based on the speaker's perception of an upcoming delay, with speakers producing *um* before longer delays and *uh* before shorter delays (Clark & Fox Tree, 2002). Various reasons may thus explain why speech is not perfectly fluent. Ultimately, speech disfluencies are regularly produced when speakers are uncertain (Smith & Clark, 1993), and can take various forms.

Moreover, the distribution of disfluencies within a given discourse is not random. For example, speech disfluencies such as *uh* are more likely to occur at the beginning of an utterance (Shriberg, 1996). They also tend to occur before a syntactic clause (e.g., [Look at *thee, uh...*[the house]], where [] delimits a clause; Clark & Wasow, 1998), and in particular, to be produced after the first word of a clause, which is often *the* in English (e.g., [the *uh*…. house]). Finally, disfluencies are regularly produced before a word that is new in the speech context (Arnold et al., 2000; Hawkins, 1971). For example, if an individual having a conversation about tea suddenly changes the topic and talks about a car, he or she is more likely to produce speech disfluencies than if continuing to talk about tea. The rationale is that when a speaker is talking about a specific topic (e.g., tea), related words and concepts are activated in memory (Meyer &

Schvaneveldt, 1971), making access to these items easier than other non-related items. Therefore, when a speaker suddenly changes the topic and introduces words that are new to a particular discourse (e.g., a car), retrieving words related to this new topic may be slightly more difficult and incur production difficulties (as reflected by *uh*s and *um*s). Together, previous research on adults' production of speech disfluencies suggests that in a sentence, a disfluency such as *uh* is likely to follow a determiner, and to precede a word that is novel (as opposed to familiar) or new (as opposed to old) in a conversation.

A final characteristic of speech disfluencies that is relevant in the context of bilingualism is that they tend to share common phonetic similarities across languages. For instance, a lengthened determiner may indicate that a speech disfluency will follow. In English, a lengthened determiner *the* – pronounced thee – often precedes silent and filled pauses and is more often used than the non-lengthened *the* – prounounced *thuh* – when anticipating a problem in word retrieval (Fox Tree & Clark, 1997). Cross-linguistic studies of speech disfluencies have revealed that many languages also make use of similar speech disfluencies, though they use different phonemes. For instance, the *uh* and *um* of English correspond to *euh* and *eum* in French (pronounced [œ] and [œm], respectively; Duez, 1982) and to *äh* and *ähm* in German (pronounced [a] and [am]; Fischer, 2000). In these three languages and others (Clark & Fox Tree, 2002), the vowel of the disfluency may be lengthened to various degrees, from a few milliseconds to several seconds. Moreover, the particular vowel used in each language may be attributable to each language having a default vowel that resembles the vowel of a common determiner in that language (e.g., English 'the' and *uh*, French 'le' and *euh*). The particular realization of speech disfluencies varies across different languages, but common characteristics include a lengthened determiner, and a lengthened vowel of the filled pause disfluency.

Most of the work detailing the nature and role of speech disfluencies in communication has been studied in the speech of adults intended to other adults, also called adult-directed speech (Soderstrom & Morgan, 2007). Some work has also investigated the presence of speech disfluencies in the speech of adults that is directed to infants, also referred to as infant-directed (or child-directed) speech. Infant-directed speech characteristically involves higher pitch, exaggerated prosody, and is slower than adult directed speech (Fernald & Simon, 1984). Because infant-directed speech is relatively slow, it tends to be more fluent than adult-directed speech, and to involve more pausing. However, these pauses are fluent prosodic breaks more often than they are disfluent ones (Soderstrom, Blossom, Foygel, & Morgan, 2008). As children get older, adults tend to be increasingly more disfluent in their child-directed speech (Nilsson Björkenstam, Wirén, & Eklund, 2013). In part because more breaks can facilitate speech processing, infants tend to prefer infant-directed speech over adult-directed speech (e.g., Cooper & Aslin, 1990). As children get older and acquire the ability to segment words from the speech stream (e.g., Jusczyk & Aslin, 1995; Jusczyk et al., 1999), they may increasingly attend to adult-directed speech, which is faster than infant-directed speech, and contains more disfluencies (Broen, 1972; Nilsson Björkenstam et al., 2013). Thus, while children initially attend more to infant-directed speech, they also hear adult-directed speech from their caregivers. A challenge that toddlers face is therefore to comprehend speech from adults that naturally contains speech disfluencies.

1.2.1 The role of speech disfluencies in comprehension

Speech disfluencies serve various purposes for speakers, but also for listeners. Speech disfluencies can also be useful for listeners who are trying to comprehending speech. As discussed in the previous section, disfluencies are frequent, and occur in a systematic fashion. Moreover, they carry a lot of information: Although speech disfluencies such as *uh* are not words per se (although see Clark & Fox Tree, 2002, who argue that they could be considered words), they convey important information. For example, they signal to listeners that a speaker is hesitating, and that the following word is likely to be novel or new to the discourse. If listeners were able to process them, they could gain highly relevant predictive information as they are listening.

Initial research on comprehension of disfluencies was conducted with adults, and suggests that disfluencies have predictive properties that are accessible to listeners (Brennan & Schober, 2001), allowing them to infer a speaker's intention and to anticipate novel words and novel referents in a discourse (Arnold et al., 2003; Arnold et al., 2007; Barr & Seyfeddinipur, 2010; Arnold, Tanenhaus, Altmann, & Fagnano, 2004). Adults' use of disfluencies as predictive cues is, however, highly based on speaker and situation characteristics. For example, adults do not use disfluencies predictively when they are told that a speaker has difficulty naming words (for example due to a clinical condition such as object agnosia), is a non-native speaker, or when they perceive a label to be difficult to pronounce (Arnold et al., 2007; Bosker, Quené, Sanders, & de Jong, 2014; Heller, Arnold, Klein, & Tanenhaus, 2015). Adult listeners are therefore flexible in their use of disfluencies as a predictive cue, and rely on the knowledge state of a speaker to decide whether to use disfluencies as a predictive cue during listening.

There has been much less work on children's processing of disfluencies. Can toddlers, like adults, also use the information inherent to speech disfluencies to guess what a speaker will refer to? This question is highly relevant, as the ability to infer a speaker's intent is an important aspect of word learning (see Tomasello, 2000), and may help children identify the correct object in an array of objects in their environment. More specifically, a disfluency such as *uh* may help children attend to discourse-new objects, and find these objects as more likely referents for the word to follow compared to familiar objects. Previous research indicates that around age 20-23 months, infants are able to distinguish fluent speech from disfluent speech (showing an attentional preference for fluent over disfluent speech; Soderstrom & Morgan, 2007). Having acquired this ability, monolingual English toddlers can use disfluencies predictively during speech processing. For example, upon hearing the English disfluency *uh*, children around age 2 look at a novel object (e.g., with a novel word-like name such as 'wug') over a familiar one (e.g., a ball; Kidd et al., 2011), unless a speaker shows evidence that they forgot the name of an object (Orena & White, 2015). In these comprehension studies with children, the novel word was both novel relative to the familiar word and new in the discourse (i.e., it had not been previously labeled in the discourse). Subsequent research has suggested that the novelty of a word within a discourse – as opposed to the novelty of a word relative to a familiar word – may be the main driver of the disfluency effect in children (Owens & Graham, 2016; Owens, Thacker, & Graham, 2017). Thus, research suggests that children – at least English-learning toddlers – can use speech disfluencies to predict that a speaker will name a novel, discourse-new object.

In addition to understanding the role of speech disfluencies and discourse context when establishing reference, children also understand that disfluencies may provide cues about a speaker's preference for certain objects. For example, children predict that a speaker who explicitly mentioned preferring objects of a certain colour (e.g., by saying "I like things that are blue") will likely prefer and label an object of that colour (i.e., blue) over another colour (e.g., pink; Thacker, Chambers, & Graham, 2018). However, when the speaker produced the disfluency *uh* before naming an object (e.g., "Look! Look at the uh… X!"), children were more likely to consider the object that was *not* the colour that the speaker preferred (i.e., pink) as a possible referent. In this case, object novelty and discourse novelty were not factors of interest:

Both objects on the display were novel and new to the discourse. In order to make predictions about which object the speaker intended to label, children had to make inferences based on both the speaker's preference and speech hesitations. Speech hesitations are thus useful for prediction during comprehension, and represent cues about a speaker's intent and preferences.

The research on the comprehension of words in disfluent speech is growing, providing more insights about the role of speech disfluencies in the context of language acquisition. However, the comprehension research with children has only been conducted with populations of monolingual children exposed to English, which is not representative of the diverse language backgrounds of other children acquiring language, such as monolinguals speaking languages other than English and bilingual children. Studying how these children comprehend words following disfluencies would provide additional insights about speech disfluencies, and whether or not they support language acquisition in children. I will return to this point in section 1.3 below.

1.2.2 The role of speech disfluencies in word learning

In development, comprehension necessarily takes place after a word has been learned. In the present research, a new question emerged when conceptualizing research on the comprehension of words following speech disfluencies: What is the role of speech disfluencies in word learning? As discussed above, disfluencies tend to be produced before novel words or words that are new to the discourse context (Arnold et al., 2000; Arnold & Tanenhaus, 2011). Because disfluencies precede novel or discourse-new words, they might affect the learning of those novel words that follow. In the following paragraphs, I will elaborate two competing predictions: that disfluencies hinder word learning, and that disfluencies help word learning.

Why might disfluencies hinder word learning? Based on the literature on the role of disfluencies in communication, disfluencies indicate uncertainty on the part of the speaker. As a result, young listeners may interpret the uncertainty as a cue that this speaker is unreliable. Previous research on the predictive use of speech disfluencies has found that children do not use disfluencies predictively when hearing a forgetful speaker compared to a knowledgeable speaker (Orena & White, 2015). Based on this research, one would predict that children may not be convinced that the speaker is using the appropriate label for a word as the disfluency might indicate that they have forgotten or are uncertain of the label for an item, and therefore children may not want to learn that word. Thus, if speech disfluencies affect children's willingness to learn new words based on a speaker's perceived lack of knowledge, disfluencies could pose a problem for word learning.

On the other hand, speech disfluencies are breaks in the flow of speech, which in turn correspond to a certain degree of pausing. This break in speech may constitute an ideal condition for word learning. First, disfluencies isolate the word to be learned, which may help children identify the word boundaries in an otherwise continuous speech stream (e.g., Brent & Cartwright, 1996; Lew-Williams, Pelucchi, & Saffran, 2011). In this respect, disfluencies may be comparable to other cues in speech that help children find word boundaries within speech (e.g., word stress; Johnson & Jusczyk, 2001), an ability that is important for language acquisition (Jusczyk & Aslin, 1995; Saffran, Aslin, & Newport, 1996). Second, if children direct their attention to the novel word upon hearing a disfluency (Kidd et al., 2011; Orena & White, 2015), this increased attention to the novel object may facilitate the novel word-object association, and in turn facilitate learning.

To date, no published study has explored the role of speech disfluencies in word learning, leaving it unclear whether they hinder or help word learning. Studying how children learn words immediately following disfluencies provides a great opportunity to gain a better understanding of how children use cues inherent to natural speech to acquire language.

1.3 Bilingualism as a Useful Lens

Most research on speech disfluencies focuses on monolingual development, and it is unclear how bilingual development fits in the picture. In this section, I make the case for bilingualism being a useful perspective from which to gain a better understanding of speech disfluencies. I begin by highlighting key differences between children raised in monolingual versus bilingual homes, and follow by highlighting similarities between the two.

Some characteristics make monolinguals and bilinguals fundamentally different, which in turn make bilingualism a useful tool to explore how disfluencies are understood and processed by young listeners. A first and obvious difference is that while monolinguals only hear and learn one language, bilinguals hear and learn two languages (Grosjean, 1982). For a child raised in a monolingual home, exposure occurs almost exclusively in one language. Therefore, all of a monolinguals' time is typically spent hearing a single language. On the other hand, for a child raised in a bilingual home, language exposure is divided between two languages. Assuming that monolinguals and bilinguals hear about the same amount of language overall, this means that monolinguals have more exposure to a particular language than bilinguals.

Even between two bilingual children, language exposure is not necessarily the same. When it comes to the amount of exposure that bilingual children have in each language, different children hear their two languages in different proportions. As such, bilinguals tend to be particularly heterogeneous as a group (Werker & Byers-Heinlein, 2008). Thus, bilinguals' exposure to one of their language will depend on relative exposure compared to their other language. Few bilingual children have truly balanced exposure to their two languages. Typically, children will hear more input and be more proficient in one language, called their dominant language. In contrast, they will hear relatively less input and be less proficient in their other language, called the non-dominant language. Language dominance matters for bilinguals. For example, 20-month-old English-French bilingual infants process speech more efficiently in their dominant language compared to their non-dominant language (Byers-Heinlein, Morin-Lessard, & Lew-Williams, 2017). This was evidenced by slower processing when there was a switch from their dominant to their non-dominant language in the same sentence (e.g.,"Find the *chien*!") compared to when there was no switch (e.g., "Find the dog!"). This is arguably because the dominant language was strongly activated during listening, making it difficult to inhibit it and switch to activating the non-dominant language. However, this slower processing as a result of a language switch was not observed when switching from the non-dominant to the dominant language (e.g., "Trouve le *dog*!"), revealing less efficient processing in and less activation of the non-dominant language. More exposure to a particular language (i.e., the dominant language) is linked to a larger vocabulary in that language compared to the other language with less exposure (i.e., the non-dominant language) (Hoff et al., 2012; Hurtado, Grüter, Marchman, & Fernald, 2014; Legacy, Zesiger, Friend, & Poulin-Dubois, 2016). Similarly, faster processing during comprehension in one language is linked with a larger productive vocabulary in that language (Marchman, Fernald, & Hurtado, 2010). This larger productive (Hurtado et al., 2014) and receptive (Hurtado et al., 2014; Legacy et al., 2016) vocabulary size is, in turn, linked to faster real-time speech comprehension. Considering in which language bilingual listeners hear speech – whether in their dominant or non-dominant language – is therefore important when studying bilinguals, as more exposure to one language is generally linked with higher proficiency in that language.

What does exposure to two languages mean in terms of speech disfluencies for a child raised in a bilingual home? Importantly in the case of speech disfluencies, exposure to two languages means that bilinguals not only hear disfluencies in two languages (e.g., English-French bilinguals hear both the English *uh* and the French *euh*), but have had different relative amounts of exposure and experience hearing disfluencies in each of their languages. Because speech disfluencies are different in form in different languages, bilingualism can serve as a useful lens into their role in language acquisition, and how they are processed by children acquiring one or two languages. For example, a question of interest is whether the English disfluency *uh* and the French disfluency *euh* will be equally useful to bilingual children as a function of their language dominance (e.g., using disfluencies more readily in their dominant language compared to their non-dominant language), and whether children will use both disfluencies to make predictions about a speaker's referential intentions. Perhaps an even stronger test of whether the use and understanding of disfluencies is tied to a listener's particular language, is whether monolingual children will use non-native speech disfluencies (e.g., an English-learning child hearing *euh*) predictively as reliably as those that they hear in their native language (e.g., the same child hearing *uh*; Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016).

Finally, it is important to emphasize that monolinguals and bilinguals do share similarities in their language acquisition experience, both generally and in the speech disfluencies that they hear. For example, bilinguals acquire their languages within monolingual children's normal range of variation (Petitto et al., 2001; Petitto & Kovelman, 2003; though see Hoff et al., 2012), and generally reach language milestones at about the same age as monolinguals (e.g., babbling, first word production; Holowka, Brosseau-Lapré, & Petitto, 2002). Bilinguals also know words for the same number of concepts as their monolinguals peers (Marchman et al., 2010; Pearson, Fernández, & Oller, 1993; Pearson & Fernández 1994), and have the same or a higher total vocabulary size than monolinguals when their two languages are combined (Core, Hoff, Rumiche, & Señor, 2013; Hoff et al., 2012; Legacy et al., 2016; Pearson et al., 1993). Thus, research supports the claim that children raised in bilingual homes can learn two languages successfully, the same way that children raised in monolingual homes do (Byers-Heinlein & Lew-Williams, 2013). When looking at the disfluencies across languages, we find that they share similarities such as lengthening of the determiner and of the vowel (Clark & Fox Tree, 2002; Fox

Tree & Clark, 1997). Thus, monolinguals and bilinguals share a similar experience in that they both hear disfluencies in the language(s) that they hear, and these disfluencies share common acoustic properties.

Overall, monolinguals and bilinguals have similar yet also different language acquisition experiences, which is also reflected in their experience with and exposure to speech disfluencies. Monolinguals' and bilinguals' similar and different experiences with language can shed light on whether or not the predictive use of speech disfluencies is based on such experiences.

1.4 Dissertation Research Objectives

As discussed above, natural speech is highly disfluent. Yet, much of the existing experimental work on language acquisition does not reflect this reality: The vast majority of experimental work with children uses fluent speech only, which is not representative of what children hear. Some work using disfluent speech has shown that monolingual English children use disfluencies to their advantage during comprehension. Work that branches out to other linguistic populations of children is crucial: Such work would shed light on how disfluencies from two languages are used by listeners from different language backgrounds as they comprehend and learn novel words. The overarching goal of this dissertation is to better understand the role of speech disfluencies in language acquisition, by considering listeners of monolingual and bilingual backgrounds. Specifically, my main prediction for this dissertation was that listeners' exposure to sounds and words specific to their native language(s) would shape language processing, which would in turn affect how speech disfluencies are processed and used during real-time comprehension and in word learning.

Chapter 2 presents research from Manuscript 1, which investigates the comprehension of words immediately following disfluencies in two studies with monolingual and bilingual 32-month-old toddlers (Study 1) and bilingual adults (Study 2). These studies help elucidate whether the exposure to disfluencies in a specific language matters for comprehension of language-specific disfluencies (e.g., English *uh* vs. French *euh*), or whether the effect of disfluencies previously observed in monolinguals is robust across listeners of different language backgrounds and expertise. Chapter 3 presents research from Manuscript 2, which extends findings from Manuscript 1, and moves beyond the comprehension aspect of disfluencies to the learning of words following disfluencies in monolingual and bilingual 32-month-old toddlers. This study helps establish the role of disfluencies in word learning. Together, the studies comprised in this dissertation contribute to the literature on the real-time processing of disfluencies, which is an integral component of word comprehension and word learning. As such, this dissertation advances the understanding of disfluencies as an information-rich component of natural speech, and their role in toddlers’ language development.

2 CHAPTER 2: Manuscript 1 on Word Comprehension

Speech is a well-structured signal, and regularities in the speech signal allow listeners to form predictions as they listen to speech in real time. For example, familiarity with English syntax allows listeners to predict that the sentence "Cats like to chase…" will likely end with the noun "mice". Listeners can also make predictions from seemingly erratic parts of speech that provide predictive information. For example, unintentional hesitation markers such as *uh* tend to appear at the start of syntactic clauses (Clark & Wasow, 1998), and to be produced immediately after the first word of a new clause, which tends to be "the" in English, for example, "Did you see the, *uh*, chameleon in the tree?". *Uh*s and *um*s also tend to precede words that are new or previously unmentioned in the discourse (Arnold, Kam, & Tanenhaus, 2007; Arnold & Tanenhaus, 2007), since speakers may be less acquainted with these new or infrequent terms (Smith & Clark, 1993). Monolingual children and adults use such speech disfluencies to predict that a speaker will refer to a new object (Arnold et al., 2007; Kidd, White, & Aslin, 2011; Orena & White, 2015). However, the basis of this prediction is not clear. The current study asks what makes speech disfluencies useful for real-time speech processing for children and adults, by investigating the processing and use of disfluencies by monolingual and bilingual listeners.

The Nature of Speech Disfluencies

Speech disfluencies provide a valuable cue to help language listeners efficiently process natural speech, as they occur regularly in predictable situations. Speakers commonly produce non-pathological stumbles that affect the flow of speech, such as filled pauses (e.g., *uh, um*), repetitions (e.g., *the the...*), and silent pauses (Fox Tree, 1995). The frequency and type of disfluencies produced vary based on factors such as the complexity of a topic (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001), and uncertainty about a topic (Smith & Clark, 1993). Disfluencies are especially likely to occur before a difficult or unfamiliar word, or one that is new in the speech context (Arnold et al., 2000; Arnold & Tanenhaus, 2007). For example, when English speakers anticipate a short delay (e.g., before producing a new or infrequent word), they often produce filled pauses such as *uh,* and when anticipating a longer delay, produce more produce more *ums* (henceforth "speech disfluencies" or "disfluencies"; Clark & Fox Tree, 2002).

The exact realization of speech disfluencies depends on the language being spoken. For example, typical English disfluencies are *uh* and *um* (Shriberg, 2001). However, typical French disfluencies are *euh* or *eum* (Duez, 1982). The primary difference between these disfluencies is in the vowel, which may be due to each language having a neutral vowel that is close to the vowel of a common determiner in a language (e.g., *uh* in English is close to "the", and *euh* in French is close to "le"). In the case of English and French, each disfluency only includes vowels present in each respective language. In other words, the vowel sound *uh* is not in the Canadian French phonemic inventory, and conversely the vowel sound *euh* is not in the Canadian English phonemic inventory. Although the realization of disfluencies differs between languages, the change in phonemes between different languages (e.g., *uh* vs. *euh*) does not change the meaning of the disfluency per se. It may merely suggest the language to which the disfluency belongs. Importantly, unlike words, speech disfluencies make no semantic contribution to an utterance.

Listeners' Use of Speech Disfluencies Across Development

Can listeners use speech disfluencies to make predictions about language? As noted above, disfluencies occur in predictable locations in the speech stream – often when a speaker is uncertain. Research with monolingual adults has demonstrated that the English filled pause *uh* helps adults recognize words on a screen faster compared to when *uh* has been edited out (i.e., cut out, but not replaced by a silent pause; Fox Tree, 2001, though see Corley & Hartsuiker, 2011). Other research with English-speaking adults has revealed that disfluent discourse leads listeners to expect that a speaker will refer to novel or unfamiliar objects (Arnold et al., 2007; Barr & Seyfeddinipur, 2010), whereas fluent discourse makes listeners expect that a speaker will continue talking about previously-mentioned objects (Arnold, Tanenhaus, Altmann, & Fagnano, 2004). Building on these findings, recent research suggests that the predictive use of disfluencies may also depend on situation- and speaker-specific characteristics such as when a speaker has difficulty naming novel words, is producing disfluencies in a non-native language, or when a label is perceived to be difficult to pronounce (Arnold et al., 2007; Bosker, Quené, Sanders, & de Jong, 2014; Heller, Arnold, Klein, & Tanenhaus, 2015). For these reasons, when listeners hear a disfluency such as *uh*, they can predict that a speaker will refer to a label that is difficult to name and therefore potentially also novel.

Research with monolingual children suggests that speech disfluencies also support language comprehension in younger learners, who use disfluencies to predict which object a speaker will name. In a study particularly pertinent to the present research, Kidd, White, and Aslin (2011) investigated 16- to 32-month-old children's real-time processing of sentences containing a speech disfluency. Monolingual English children viewed object pairs on an eyetracker. Half of the objects were novel objects with nonsense labels (e.g., mog), while the other half were familiar objects with familiar labels (e.g., shoe). On each trial, a familiar-novel object pair was presented three times. On the first two presentations, the familiar object was named to establish its familiarity in the discourse (e.g., "I see the *shoe*! Ooh, what a nice *shoe*!"). On the third presentation, either the familiar or novel target word was named in a fluent sentence (e.g., "Look! Look at the shoe/mog!") or a disfluent sentence (e.g., "Look! Look at *thee uh* shoe/mog!"). Children's looking at the two objects was measured from the onset of the disfluency until just before the speaker uttered a label (i.e., the disfluency period). Children aged 28–32 months reliably looked predictively towards the novel object when they heard the disfluency. However, children aged 16–20 months did not appear to use the disfluency to predict the speaker's intended target, suggesting that the ability to use disfluencies predictively appears over the course of development.

Subsequent research has extended this work to show that children's predictive use of speech disfluencies is robust with previously knowledgeable speakers, but not with previously forgetful speakers (Orena & White, 2015). Additionally, 3-year-olds can use disfluencies to predict that a speaker will label an object that is perceptually familiar but novel to the discourse (e.g., a familiar object that has not previously been mentioned; Owens & Graham, 2016). However, while adults can use filled pauses to anticipate reference solely based on object novelty, 3- and 5-year-olds require that an object be new in the discourse to anticipate reference to a novel object (Owens, Thacker, & Graham, 2017). Together, findings from real-time comprehension studies with monolingual children suggest that their ability to use speech disfluencies predictively emerges around their second birthday, and that they flexibly adapt their predictions based on both speaker and context.

Speech Disfluencies as Informative Cues

Important developmental questions remain about what makes speech disfluencies informative to listeners. One possibility is that children learn to use disfluencies predictively through experience with specific linguistic forms: they learn that *uh* and *um* often precede the labeling of novel or discourse-new referents. Both children and adults are highly sensitive to the statistical regularities present in human language, and they use these regularities for language acquisition and processing. For example, young infants are sensitive to when particular syllables co-occur or pattern in specific ways (Gómez & Gerken, 2000; Saffran, Aslin, & Newport, 1996).

Experience with a language's patterns of co-occurrence allows listeners to gain knowledge of the structure and constraints of their language.

Knowledge about co-occurrence patterns in language can be used predictively during comprehension. For example, both children and adults predict that a speaker is more likely to refer to a food than a non-food item if they have heard the verb "eat" (e.g., Friedrich & Friederici, 2005; Snedeker & Trueswell, 2004). In the context of speech disfluencies, listeners could leverage their repeated experience hearing specific disfluencies prior to novel or discourse-new words to form predictions that a disfluency is likely to precede a novel or discourse-new label. It may be that children become able to use disfluencies predictively once they accrue sufficient experience that disfluencies such as *uh* tend to precede a speaker's labeling of a novel object. Moreover, this account would predict that children would be best able to use typical, frequently-encountered disfluencies as a basis to make predictions about what will follow in the sentence, but would be less able to use atypical, non-native disfluencies.

Some evidence from the adult literature supports this account, showing that the form of the speech disfluency could affect its use in speech prediction. In a study that compared adult's predictions in the context of disfluencies produced by either a native versus an accented speaker, adults increased attention to low-frequency referents (e.g., a sewing machine) only when hearing the native speaker (Bosker et al., 2014). The main interpretation of this result was that listeners attributed the disfluency to retrieval difficulty in the native speaker, but not in the accented speaker. However, an alternate possibility is that the non-native realization of the disfluency itself disrupted adults' ability to use the disfluency predictively. In support of this idea, over the course of the study, listeners in the native condition increased looking to the low-frequency referent upon hearing the disfluency, suggesting that adults detected and learned the association between disfluencies and subsequently labeled referents. However, as both the disfluency and the rest of the sentence were accented, this study does not provide conclusive evidence of whether atypical pronunciation of the disfluency itself disrupted its predictive use. Nonetheless, if listeners' predictions are less robust when hearing a non-native disfluency, this would suggest that such predictions are, at least in part, driven by experience hearing the specific native realization of a disfluency prior to a speaker naming novel or discourse-new referents.

In contrast, it is also possible that listeners can flexibly use a range of speech disfluencies, whether native or non-native, to predict that a speaker will label a novel object. Although languages differ in the typical realizations of disfluencies, disfluencies often involve lengthening (Fox Tree & Clark, 1997). Further, even within a single language, there can be a variety of typical disfluencies, such as *uh*, *um*, and *hmm* in English. Moreover, whatever their form, disfluencies typically occur because a speaker is uncertain about the choice of a word, or is having retrieval difficulties (Beattie & Butterworth, 1979). Thus, it is also possible that children learn that lengthening in general, either by itself or as a marker of speaker uncertainty, can provide a cue about what a speaker is likely to label. In this case, listeners might be able to make equal use of native and non-native disfluencies.

Speech Disfluencies and Bilingualism

Comparisons between monolinguals and bilinguals can help illuminate the role of language experience in children's predictive use of disfluencies. This is because bilingual individuals are unique in their experience with disfluencies. A first and obvious distinction between monolingual and bilingual environments is that while monolinguals hear only one language, bilinguals hear two. As such, English monolinguals are only exposed to disfluencies specific to the English language, such as *uh*, but do not encounter *euh* that is typical in French. Likewise, French monolinguals are only exposed to the French *euh,* but not the English *uh*. In contrast, English-French bilinguals hear both *uh* and *euh*.

Second, as a consequence of dividing their time between two languages, bilinguals hear a similar total number of disfluencies as monolinguals, but on average only half as many in each particular language. Further, bilinguals can vary substantially in their relative exposure to each language, and thus may have more experience with disfluencies in one of their languages (typically the dominant or most-often-heard language) than in the other (the non-dominant or least-heard language).

Finally, bilingual communities often mix their languages. When language mixing or "code-switching" (Poplack, 1980), bilinguals often speak one language while borrowing single words from another language. Speech disfluencies may also be realized in a code-switched manner. For example, bilingual adults whose native language is English may pronounce English *uh*s when speaking French, a less-proficient language. This means that bilinguals likely also hear code-switched disfluencies, or disfluencies for which the language of the disfluency and the language of the sentence do not match, in addition to hearing sentence-matching disfluencies. Thus, some disfluencies produced and heard by bilinguals may be language-consistent (e.g., the English *uh* in an English sentence, as in “Look at the *uh*… dog!”), whereas others may be language-inconsistent (e.g., the English *uh* in a French sentence, as in “Regarde le *uh*... chien!”).

These three distinctions between monolingual and bilingual experiences highlight the fact that the two language groups have different experiences with speech disfluencies. If language-specific experience with particular phonetic forms is important for the development of the predictive use of disfluencies, then bilinguals might develop this ability later than monolinguals as they have less exposure to disfluencies in each language (but likely hear as many total disfluencies across their two languages). Another prediction is that bilinguals might more readily leverage disfluencies in their dominant than in their non-dominant language, as they might hear these more frequently. Finally, because bilinguals likely encounter code-switched disfluencies, they might better be able to make use of these than monolinguals. On the other hand, we might also find that monolinguals and bilinguals are similar in when they begin to use disfluencies for prediction, and that bilinguals use disfluencies equally in their dominant versus non-dominant language. Such findings would be evidence against the idea that children begin to use disfluencies when they learn that specific syllables – such as *uh* or *um* – tend to precede novel or discourse-new labels.

The Present Research

The main goal of this research was to investigate the mechanism underlying the ability to use disfluencies to predict the content of upcoming speech. To do so, our study compared listeners from two different language backgrounds, who diverge in their experience with the specific realization of speech disfluencies: monolinguals and bilinguals. We also leveraged the fact that languages differ in how they realize disfluencies, and that listeners’ experience is limited to the type of disfluencies present in their native language(s). Finally, we took a developmental perspective, by investigating both children (aged 32 months: Study 1) and adults (Study 2).

Our experimental approach expanded on Kidd, White, and Aslin’s (2011) original study with monolingual English children. We tested monolingual and bilingual 32-month-olds and bilingual adults using a preferential-looking paradigm. This age was chosen because monolinguals are consistently able to use disfluencies predictively at this age (Kidd et al., 2011). English and French monolingual children were tested on their predictive use of two types of disfluencies: those consistent with their native language and those inconsistent with their native language. Having these two groups of monolinguals allowed us to replicate Kidd et al. (2011)'s findings with monolingual English children, and to test the generalizability of their results with monolingual French children. Importantly, it also allowed us to test whether monolinguals' predictive use of disfluencies is limited to those produced in a native-like way. We also tested English-French bilingual children and adults on their comprehension of language-consistent and language-inconsistent disfluencies, in both their dominant and non-dominant languages. Findings from these two bilingual groups are key in helping us understand how speech disfluencies – *uh* and *euh* in particular – are processed, and whether their predictive use is tied to a specific language.

We hypothesized that experience with the specific syllables that are characteristic of a language's disfluencies allow listeners to form predictions about the speech follows. As such, listeners will be especially sensitive to the predictive nature of the disfluencies that are most common in their everyday language environments. Monolinguals will be best able to use native-language disfluencies, but will be less able to use disfluencies that are inconsistent with their native language. Bilinguals will be best able to use the disfluencies that are characteristic of their dominant language, and will show somewhat less ability to leverage the disfluencies characteristic of their non-dominant language. Moreover, because bilinguals have less experience with each language than monolinguals have with their native language, their ability to use disfluencies predictively may emerge later, particularly for language-consistent disfluencies. Finally, since bilinguals have some experience with code-switched disfluencies, they might perform better on language-inconsistent disfluencies than monolinguals. In sum, we predicted that children's predictive use of disfluencies would differ depending on the specific realization of the disfluency (language-consistent vs. language-inconsistent), their own language background (monolingual vs. bilingual), and for bilinguals, whether they were hearing a disfluency from their dominant or non-dominant language.

2.1 Study 1: Children

2.1.1 Method

The present research was approved by the Human Research Ethics Committee at Concordia University, and all children were treated in accordance with the Tri-Council Policy

Statement: Ethical Conduct for Research Involving Humans (TCPS 2; CIHR, NSERC, & SSHRC, 2014). Parents provided informed consent prior to the study (see Appendix A).

Participants. Forty-eight 32-month-old children were included in the final sample. Children were recruited from government birth lists, and lived in Montréal, Canada. All children were healthy and typically-developing, with no reported hearing or vision problems. Following Kidd, White, & Aslin (2011), we aimed for a sample of 16 participants per group. There were 16 monolingual English children (M_{age}= 32.18, *SD* = 17 days, range = 31.29 to 33.20, females = 10), 16 monolingual French children (M_{age} = 32.18, *SD* = 15 days, range = 31.26 to 33.9, females = 6), and 16 English-French bilingual children (M_{age} = 32.22, *SD* = 13 days, range = 31.27 to 33.7, females = 7). Of the sixteen bilinguals, 7 were English-dominant (*M* = 65.3% English exposure, *SD* = 5.2%) and 9 were French-dominant (*M* = 62.1% French exposure, *SD* = 8.9%). Monolingual children were exposed to either English or French at least 90% of the time. English-French bilingual children were exposed to each of their languages a minimum of 25%, and did not have more than 10% exposure to a third language. All bilinguals had acquired their two languages from birth except for three children, who acquired their second language upon entering daycare at 10, 12, and 14 months, respectively. For bilingual children, language dominance was established based on exposure, with the language most-often heard considered the dominant language. Another 29 children were tested but not included in the final sample due to failure to meet pre-established language criteria for monolingualism or bilingualism (15), low birth weight or premature birth (4), fussiness/inattention (5), reported health issues (2), parental interference (2), or equipment error (1). Data collection for this study began in June 2015, and was completed in June 2017.

Measures. The Multilingual Approach of Parent Language Estimate (MAPLE) and an adaptation of the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., *under review*), were used to assess input from different caregivers across the child's life, and to measure monolinguals' and bilinguals' exposure to their language(s) (see Appendix B). In a semi-structured interview, the primary caregiver was asked about the family's language background, the child's home environment, daily activities, and typical daily routines. This interview yielded an estimate of the total number of hours spent in each language over the course of the child's life, which was converted to a percentage, and was then averaged with the parent's estimate to yield final exposure percentages.

Finally, a demographics questionnaire gathered information about the family's background and the child's health history (see Appendix C).

Stimuli. Visual stimuli consisted of 32 familiar-novel object pairs appearing against a white background. The familiar objects were imageable words from MCDI vocabulary norms in American English (Fenson et al., 1993) and Québec French (Boudreault, Cabirol, Trudeau, Poulin-Dubois, & Sutton, 2007). Based on MCDI norms, these objects were known by at least 50% of 18-month-olds in English, and by at least 50% of 16-month-olds in French (French norms for 18-month-olds were not available). Images of familiar objects were selected from Google, and images of novel objects were selected from the NOUN-2 Database (Horst & Hout, 2016).

Auditory stimuli were recorded by a native bilingual female speaker of Canadian English and Québec French with no noticeable accent in either language, who produced stimuli in child-directed speech. Auditory stimuli were normalized to a comfortable hearing level of 70dB. Novel words were chosen to sound like possible words in English or French. Names for novel words in English were selected from the NOUN-2 Database (Horst & Hout, 2016), from Kidd, White, and Aslin (2011), or were created for this study. Novel English and French words had the same number of syllables (see Table 1).

Familiar and novel objects were paired for presentation on each trial, and were matched for image size and color salience. Auditory stimuli in each pair were phonologically distinct in place of articulation at word onset, did not rhyme, were matched on number of syllables, and in French were matched for grammatical gender (Lew-Williams & Fernald, 2007). Half of French trials presented feminine words (e.g., *la* maison and *la* télue), and the other half masculine words (e.g., *le* soulier and *le* pafli). Audio-visual stimuli are available on the Open Science Framework at https://osf.io/qn6px/, and visual stimuli are displayed in Appendix E.

The sequence of each trial followed Kidd et al. (2011), such that novel labels were also discourse-new (see Owens & Graham, 2016, and Owens, Thacker, & Graham, 2017, for evidence that discourse novelty is a driver of disfluency effects in children). There were three presentations of each object label. On the first presentation, the familiar object was named in an English or a French sentence (e.g., English: "Look at the *doll*!") as the two images appeared. A black screen then appeared for 1 second. On the second presentation, the same familiar-novel object pair appeared and was named in English or French, in a different sentence (e.g., English:

“Ooh! What a nice *doll*!”). This was again followed by a black screen for 2.5 seconds. On the third presentation, the familiar-novel object pair was shown again. On half of the trials the familiar object was named, and on the other half the novel object was named. Crucially, in this third sentence, half of the trials (8) were fluent, and half (8) were disfluent. On Fluent trials, the final sentence contained no disfluency. Of the disfluent trials, half (4) were Disfluent Language-consistent, where the final sentence had an elongated determiner (thee/leee/laaa) and a filled pause consistent with the language of the sentence (e.g, English: “Look at thee uh *doll*/*rel*!”). The other half of disfluent trials (4) were Disfluent Language-inconsistent, where the final sentence contained an elongated determiner and a filled pause inconsistent with the language of the sentence (e.g, English: “Look at thee euh *doll*/*rel*!”). As in Kidd, White, and Aslin (2011), “the” was elongated for disfluent trials (pronounced “*thee”* instead of “*thuh*”) given that this is typical of determiners preceding disfluencies in natural speech (Fox Tree & Clark, 1997). Since the realization of determiners in both English and French can change when preceding a vowel, all words started with consonants. Each trial lasted 18.5 seconds.

A total of eight study orders were created (4 in English, 4 in French) that counterbalanced the side of presentation of each object, and which object was labeled on each trial (see Figure 1). All sentences within an order were in the same language. Trials were presented quasi-randomly, such that no more than two trials of the same type (Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) appeared in a row. One of five animated attention-grabbing objects accompanied by music were presented in the center of the screen between each trial to recapture the child’s attention, and to direct their gaze to the center of the screen. Monolingual children were tested in only one order in their native language (randomly-assigned), and saw a total of 16 trials. Bilinguals were tested in two orders (one randomly-assigned English and one French) and saw a total of 32 trials. For bilinguals, the order of presentation of English and French was counterbalanced.

Procedure. A Tobii T60XL eyetracker was used to present the stimuli on a 24-inch screen, while eye-gaze data were gathered at a sampling rate of 60 Hz (Tobii Eye Tracker, 2009). Children sat on their parent’s lap or by themselves on a chair in a dimly-lit soundproof room, approximately 60 cm away from the screen. The caregiver was instructed not to interact with their child during the study, and wore opaque sunglasses and headphones through which music was played to ensure that they were blind to the experimental condition. The study began following a 5-point infant calibration routine to calibrate the eyetracker to children's eyes. A play break was given to bilingual children between the two experimental orders if needed. The total duration of the experiment was approximately 6 min for monolinguals and 12 min for bilinguals. Questionnaires were completed following the study.

2.1.2 Results

Analytic Strategy. Consistent with Kidd and colleagues' study (2011), the crucial part of the present study was to assess looking during the disfluency period on disfluent trials, compared to an equivalent window on fluent trials. This was a 2-second window before the onset of the target word in the third sentence of each trial, during which participants heard the speech disfluency. The target word always occurred 15 s into each trial (4000 ms into sentence 3), meaning that the Disfluency Window was from 13 to 15 s into each trial (2000–4000 ms into sentence 3). The window was shifted by 250 ms as in Kidd, White, and Aslin (2011) to allow for time needed for children to launch an eye movement. This window (2250–4250 ms) establishes the predictive nature of disfluencies; if children look to the novel object as they hear a disfluency, and importantly, *before* they hear a target word, this means that they effectively used disfluencies to predict that the speaker was going to label a novel object. In addition to the main question of children's looking during the disfluency period, secondary analyses were conducted to assess looking to the target object (vs. the distractor) after it was named in the Label Window, to examine children's processing of the familiar and novel words.

Eye-gaze data collected was summed across time for each trial and within each area of interest using the R package eyetrackingR (Dink & Ferguson, 2015; R Core Team, 2017). On each trial, rectangular areas of interest were established about 2 cm around each object. The proportion of looking to the novel object, obtained by dividing the total duration of fixation at the novel object divided by the sum of the total fixation duration of fixation at both the novel and familiar objects, was the main dependent variable. Analyses were performed separately for monolingual and bilingual children, as monolinguals completed the study only in their native language, while bilinguals completed the study in their two languages.

Monolinguals.

Disfluency Window. Average proportion of looking time data are displayed in Figure 2. Preliminary analyses revealed no differences as a function of gender or between the monolingual groups (English monolinguals, French monolinguals), therefore data were collapsed over these factors.

We conducted a one-way repeated-measures ANOVA with Trial type (3 levels: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) as the independent variable, and proportion of looking to the novel object as the dependent variable. Results showed a significant main effect of Trial type, [$F(2, 62) = 4.99$, $p = .01$, $\eta^2 = .073$]. Follow-up paired-samples *t*-tests revealed that monolinguals looked less to the novel object on Fluent trials ($M = .51$, $SD = .17$) than on Disfluent Language-consistent trials [$M = .61$, $SD = .21$, $t(31) = –2.523$, $p = .02$, $d = 0.52$], and on Disfluent Language-inconsistent trials [$M = .62$, $SD = .20$, $t(31) = –2.534$, $p = .017$, $d = 0.59$]. More central to the current research question, there was no difference between Disfluent Language-consistent ($M = .61$, $SD = .21$) and Disfluent Language-inconsistent ($M = .62$, $SD = .20$) trials, [$t(31) = –0.310$, $p = .76$, $d = 0.05$]. Results of follow-up *t*-tests compared to chance are presented in Table 2. Monolinguals' proportion of looking to the novel object was above chance on both Disfluent Language-consistent trials and on Disfluent Language-inconsistent trials, but not on Fluent trials. Together, these results replicate previous findings that monolinguals of this age use disfluencies predictively while replicating this finding in a new group (French monolinguals), and demonstrate that monolinguals are flexible in this capacity across the specific realization of disfluencies.

Label Window. See Figure 3 for a timecourse of average looking to the target. To ensure that monolinguals looked at the target object once it was named, we analyzed looking in a 2 s window anchored on the target word onset (4000 ms into sentence 3) on each trial type (Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent). As with the Disfluency Window, the Label Window was shifted by 250 ms to allow for the launching of eye movements (4250–6250ms). We also assessed whether it was more difficult for monolinguals to identify novel targets compared to familiar targets. A 3 (Trial type: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) X 2 (Target type: Familiar, Novel) repeated-measures ANOVA on the proportion of looking to the target revealed no significant main effect of Trial type, [$F(2, 59) = 0.10$, $p = .91$, $\eta^2 = .082$], and no main effect of Target type, [$F(1, 59) = 0.72$, $p = .41$, $\eta^2 = .613$]. There was also no significant interaction between Trial type and Target type, [$F(2, 350) = 1.81$, $p = .17$, $\eta^2 = 1.805$]. These results suggest that looking to the target once it was labeled for monolingual children did not differ based on different trial types, or based on whether the target was novel or familiar. Moreover, children looked at the labeled target significantly above chance on all trial types (*p*s $< .05$), suggesting that they could equally identify the target on familiar and novel label trials.

Bilinguals.

For bilinguals, we conducted parallel analyses to monolinguals, but also investigated the effect of language dominance on the predictive use of disfluencies. One bilingual participant was inattentive during testing in the non-dominant language, and was only retained for analyses pertaining to the dominant language.

Disfluency Window. Results for bilinguals are displayed in Figure 2. Preliminary analyses revealed no effect of gender, and subsequent analyses were collapsed over this factor. A 3 (Trial type: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) X 2 (Language of testing: Dominant language, Non-dominant language) repeated-measures ANOVA on the proportion of looking to the novel object revealed a significant main effect of Trial type, [$F(2, 57) = 8.27$, $p = .001$, $\eta^2 = .290$]. There was no effect of Language of testing, [$F(1, 25) = 0.03$, $p = .86$, $\eta^2 = .001$], and no interaction between Trial type and Language of testing, [$F(2, 57) = 1.78$, $p = .18$, $\eta^2 = .062$]. These results suggest that language dominance did not affect looking to the novel object for bilinguals, thus this factor was collapsed across subsequent analyses.

T-tests were used to follow up on the main effect of trial type. Bilinguals showed a similar pattern to monolinguals. They looked significantly less to the novel object on Fluent trials ($M = .50$, $SD = .18$) than on Disfluent Language-consistent trials [$M = .66$, $SD = .23$, $t(30) = –3.274$, $p = .003$, $d = 0.77$], or on Disfluent Language-inconsistent trials [$M = .65$, $SD = .26$, $t(29) = –2.859$, $p = .008$, $d = 0.67$]. As with monolinguals, there was no difference in looking for bilinguals between Disfluent Language-consistent ($M = .66$, $SD = .23$) and Disfluent Language-inconsistent trials [$M = .65$, $SD = .26$, $t(29) = 0.254$, $p = .80$, $d = 0.04$]. Results of follow-up *t*-tests compared to chance are presented in Table 2. Looking to the novel object was statistically above chance on both Disfluent Language-consistent trials, and on Disfluent Language-inconsistent trials, but not on Fluent trials. These comparisons suggest that bilingual children are not only successful at using disfluencies to predict novelty, but that they are equally successful when disfluencies were consistent and inconsistent with the language of the sentence.

Label Window. Lastly, we assessed whether bilinguals looked to the target object after it was labeled (see Figure 3 for a timecourse of average looking results). We analyzed looking in the 2 s after the target object was labeled. A 3 (Trial type: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) X 2 (Target type: Familiar, Novel) repeated-measures ANOVA on the proportion of looking to the target revealed no main effect of Trial type, [$F(2, 25) = 0.89$, $p = .42$, $\eta^2 = .072$]. However, there was a marginally-significant main effect of Target type, [$F(1, 57) = 2.98$, $p = .09$, $\eta^2 = .052$], suggesting that bilinguals may not look to novel targets as much compared to familiar targets. Finally, there was no significant interaction between Trial type and Target type, [$F(2, 56) = 1.24$, $p = .30$, $\eta^2 = .043$]. These results suggest that bilingual children's looking to the target differed when the target was familiar versus novel, but not when trials were fluent or disfluent.

We further investigated looking to familiar versus novel targets using *t*-tests. Overall, bilingual children spent more time looking at the target when its label was familiar ($M = .71$, $SD = .27$) than when it was novel [$M = .61$, $SD = .23$, $t(30) = 2.26$, $p = .03$, $d = 0.41$]. However, looking to both familiar and novel targets was significantly above chance (.50), [$t(30) = 17.582$, $p < .001$, $d = 6.42$], and [$t(30) = 39.397$, $p < .001$, $d = 14.38$], respectively, and this result held when looking at this factor by Trial type (*p*s < .05). Note that success at novel label trials requires the use of a disambiguation heuristic such as mutual exclusivity, a point we will return to in the discussion. These findings suggest that bilingual children could identify the referent of both familiar and novel labels, but that their performance was less robust for novel labels.

2.1.3 Discussion

Study 1 tested monolingual and bilingual 32-month-olds on their responses to disfluencies that were language-consistent and language-inconsistent. All children successfully used disfluencies to predict novelty, whether or not the language of the disfluencies was consistent with the language of the sentence. Performance upon hearing the disfluency was similar for both monolinguals and bilinguals, and across bilinguals' dominant and non-dominant language. Contrary to our hypothesis, there was no effect of any of our manipulations. We did not find evidence that children benefit from experience with specific realizations of disfluencies.

While our main interest was in children's performance during the disfluency period, we also examined their looking towards the labelled target object. Importantly, all children looked at the target object above chance. For monolinguals, looking was similar for familiar and novel labels. However, bilinguals' looking was less robust during novel label trials. Indeed, a direct comparison revealed a significant difference between monolinguals and bilinguals on novel label trials [$t = 3.590$, $p < .001$] but not on familiar label trials [$t = 0.179$, $p = .51$]. In our task, children had to infer the meaning of the novel label using a heuristic such as mutual exclusivity (i.e., knowing that objects typically only have one label; Markman & Wachtel, 1988). Our findings are in line with previous reports that bilinguals are less robust in their use of mutual exclusivity than monolinguals, presumably because bilingual children learn two labels for each object (e.g., Byers-Heinlein & Werker, 2009; 2013; Houston-Price, Caloghiris, & Raviglione, 2010). Yet, this monolingual-bilingual difference has not been found in all studies (Byers-Heinlein, Chen, & Xu, 2014; Frank & Poulin-Dubois, 2002). We note that while many infant studies have used looking paradigms and found monolingual-bilingual differences, looking-based paradigms have only rarely been used to compare monolinguals and bilinguals older than age 2. Our current results raise the possibility that looking-based measures might be a more sensitive way for future studies to examine this question.

One possible alternate explanation of our disfluencies results is that children were leveraging their specific experience with disfluencies, but that the difference between *uh* and *euh* was too subtle for children to discriminate between the two. In this case, language-consistent and language-inconsistent disfluencies would have been perceived identically, and that is why children showed no difference in their performance. This explanation is somewhat unlikely, since previous research suggests that children of our participants' age and much younger can differentiate small speech sound contrasts (e.g., Byers-Heinlein & Fennell, 2014). Nonetheless, a discrimination task is needed to rule out the possibility that children could not perceive *uh* and *euh* as two different sounds. However, while it would be ideal to test 2.5 year-old children on their discrimination of the English and French disfluencies, this is methodologically challenging. Speech sound discrimination is typically tested in infants using habituation or head-turn paradigms (Jusczyk & Aslin, 1995), both of which are inappropriate for preschool-aged children (Johnson & Zamuner, 2010; Werker, Polka, & Pegg, 1997). Older children and adults can be tested with more explicit tasks that use verbal responses or button presses, but 2.5 year-olds are young to perform reliably on such tasks. Given these challenges, we addressed this issue with data from adults. In Study 2, we tested adults on the same comprehension task on which children were tested in Study 1, and in addition, adults completed a discrimination task to assess whether they could perceive the difference between English and French disfluencies.

2.2 Study 2: Adults

Study 2 set out to compare the processing of disfluencies in children to their processing by adults, who possess an expert language system. Given the lack of difference between monolinguals and bilinguals in Study 1, and the difficulty of recruiting monolingual participants in Montréal (where most adults interact regularly with both native French and native English speakers), we focused on bilingual adults. Adults were tested on the same comprehension task given to children in Study 1, as well as a discrimination task to rule out the possibility that the difference between the sounds *uh* and *euh* was too subtle to be readily perceived.

2.2.1 Method

As in Study 1, research in this study followed institutional and national guidelines for research with human participants. Adult participants read and signed a consent form (see Appendix A).

Participants. Sixteen bilingual English-French adult participants (females = 15; 14 English-dominant, 2 French-dominant) were included in the study, and ranged in age from 19 to 43 years old, with one not reporting (M_{age} = 23y). All were students at Concordia University, an English-speaking institution in Montréal, a city in which both English and French are used in the day-to-day lives of residents.

Participants filled out the Language Background Questionnaire (Segalowitz, 2009), which provided self-report data on participants' language background and proficiency (see Appendix D). This was also used to determine participants' eligibility in the study. To be included in the study, adults' native language had to be either English, French, or both. Language criteria for inclusion also required that participants have scores of at least 4/5 on self-rated comprehension and production in both French and English, with 1/5 or less on any other languages spoken. Adult participants' English comprehension scores averaged 4.9/5, while English production scores averaged 4.9/5. French comprehension scores averaged 4.7/5, and French production scores averaged to 4.3/5. For 12 adults, their dominant language was also the language that they acquired first. All adults reported having acquired their second language before the age of 6.5 years old (M_{age}= 3.8y). An additional 6 adults were tested but not included in the final sample for having a native language other than French or English (4), for falling asleep during the study (1), or for difficulty calibrating the participant's eyes to the eyetracker (1). Adults were recruited via a participant pool for Psychology students.

Comprehension and discrimination tasks. For the comprehension task, the equipment and stimuli used in the present study were identical as in Study 1. Adults sat on a chair, and were instructed to attend to the screen. After the comprehension task, participants completed a discrimination task by listening to free-field auditory stimuli on a computer. Participants listened to individual *uh* and *euh* tokens that were excised from sentences in the comprehension task, and were asked to circle on a piece of paper whether they heard an English or a French disfluency (see Appendix F). Participants had 8 s to listen to each individual disfluency and to circle their answer before the next sound was played, with no possibility of playing a sound twice. There were a total of 32 disfluencies played (16 English, 16 French) presented in a quasi-random order. The discrimination task was about 4 min in duration.

Procedure. Adults first filled out language background questionnaires to assess their language proficiency, then completed the comprehension task. The study began after the same calibration routine used with children. As with the bilingual children, adults then watched two experimental orders, one in their dominant language and one in their non-dominant language (counterbalanced). Next, adults completed the discrimination task, before being debriefed at the end of the study.

2.2.2 Results

Analytic Strategy. For the eye-tracking task, the analytic strategy for the comprehension task followed that used in Study 1. For the discrimination task, we summed the number of trials on which participants correctly identified the language of the disfluency, which yielded a maximum score of 32 if adults identified all tokens correctly.

Disfluency Window. See Figure 2 for an illustrated depiction of the proportion of looking averages. As only 1 male participated in the study, gender was not included in the analyses. A 2 (Language of testing: Dominant language, Non-dominant language) X 3 (Trial type: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) repeated-measures ANOVA conducted on the proportion of looking to the novel object showed a main effect of Trial type, [$F(2, 60) = 8.89$, $p < .001$, $\eta^2 = .296$]. However, there was no effect of Language of testing, [$F(1, 26) = 0.13$, $p = .72$, $\eta^2 = .005$], and no interaction between Trial type and Language of testing, [$F(2, 60) = 0.62$, $p = .54$, $\eta^2 = .021$]. These results suggest that language dominance did not affect looking to the novel object for bilingual adults, but that there was a main effect of Trial type as in Study 1. As was observed in the children's results, there was no significant difference in looking between the Disfluent Language-consistent (M = .30, SD = .26) and Disfluent Language-inconsistent trials [M = .30, SD = .26, $t(31) = -0.064$, p = .95, d = 0.04]. However, participants looked at the novel object significantly less on Fluent trials (M = .17, SD = .17) than on Disfluent Language-consistent trials, [$t(31) = -3.2265$, p = .003, d = 0.59], or on Disfluent Language-inconsistent trials [$t(31) = -3.73$, p = .001, d = 0.59]. Results compared to chance are displayed in Table 2. Surprisingly, comparisons against chance revealed a somewhat different pattern from the children, in that participants looked towards the novel object significantly *below* chance across all three trial types: Disfluent Language-consistent trials, Disfluent Language-inconsistent trials and Fluent trials. Together, these results indicate that, although looking to the novel object compared to the familiar object was low in general compared to children, adults' looking to the novel object increased on disfluent trials compared to on fluent trials.

Label Window. See Figure 3 for an illustrated depiction of the looking-to-target results of adults over the course of a trial. To ensure that adults were completing the main task, we examined adults' looking to the target once it was labeled. Across the three trial types, adults' looking at the target during the Label Window was significantly above chance: Fluent [M = .85, SD = .07, t = 27.267, $p < .001$], Disfluent Language-consistent [M = .84, SD = .11, t = 17.266, $p < .001$], and Disfluent Language-inconsistent [M = .85, SD = .09, t = 21.647, $p < .001$]. Thus, irrespective of hearing fluent versus disfluent speech, adults looked at the target after hearing it labelled.

Lastly, we also investigated whether looking to the target was more difficult for adults when words were novel compared to familiar. A 3 (Trial type: Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent) X 2 (Target type: Familiar, Novel) repeated-measures ANOVA revealed no main effect of Trial type [$F(2, 59) = 0.19$, p = .83, η^2 = .006], and no interaction between Trial type and Target type, [$F(2, 59) = 1.54$, p = .22, η^2 = .052]. However, there was a significant main effect of Target type, [$F(1, 374) = 81.97$, $p < .001$, η^2 = .219], indicating that adults more accurately looked at familiar targets compared to novel targets. These results suggest that although accuracy for both target types was very high, adults were more accurate at identifying the target word when it was familiar compared to novel, a finding reminiscent of results with bilingual children.

Discrimination task. On average, participants identified 97.7% of speech disfluencies correctly as belonging to either English or French, or 31.3 out of 32 (range = 24 to 32, median = 32). Thus, all participants could clearly tell apart English from French disfluencies, and obtained very high discrimination scores, well-above what would be expected from chance (50% or 16 correct answers), $t(15) = 30.25$, $p < .001$.

2.2.3 Discussion

In Study 2, we tested bilingual English-French adults on their predictive use of disfluencies. Like children, adults looked towards the novel object more when hearing disfluent compared to fluent speech. This replicates previous findings that monolinguals are sensitive to the predictive nature of disfluencies, and extends this phenomenon to a new population: bilingual adults. Moreover, this pattern of results was seen whether the disfluencies were consistent or inconsistent in language with the language of the sentence. Importantly, results from the discrimination task indicated that adults could easily discriminate the English *uh* and the French *euh,* suggesting that these were perceptually distinct. Combined with evidence that bilinguals are sensitive to a range of vowel contrasts at a far younger age than the 32-month-olds tested in Study 1 (e.g., Byers-Heinlein & Fennell, 2014), this suggests that bilinguals' similar performance in the Language-consistent and Language-inconsistent conditions was likely not due to a failure to detect the language mismatch. Given that adults detected but were not influenced by the specific realization of the disfluency (*uh* vs. *euh*), these results also fail to support our hypothesis that listeners associate particular disfluency realizations with novelty.

A second intriguing finding from this study was that after the object was labeled, adults performed better on familiar label trials than on novel label trials. In Study 1, this pattern was also seen with bilingual children, although not with monolingual children. This finding is in line with studies observing less consistent use of mutual exclusivity in bilingual children than in monolingual children (e.g., Byers-Heinlein & Werker, 2009; 2013; Houston-Price et al., 2010), and raises the possibility that differences in the use of mutual exclusivity-related behaviors such as disambiguation are affected by bilingualism into adulthood. However, the current study did not have a monolingual adult control group, and this possibility remains a hypothesis to be tested directly.

One unexpected difference between adults' and children's performance was the total amount of attention directed to the novel vs. familiar object during the disfluency period. While children tended to look either at the novel object (disfluent trials) or equally at both objects (fluent trials), adults showed a pattern of much greater overall attention to the familiar object across all trial types. Looking to the novel object increased when hearing a disfluency, but nevertheless greater overall attention was directed to the familiar object that had been labeled earlier in the sentence. This differs from previous studies, wherein adults spent more time looking at previously-unmentioned (Arnold, Fagnano & Tanenhaus, 2003; Arnold et al., 2004) and unfamiliar (Arnold et al., 2007) objects when hearing a disfluency. One important difference is that previous studies with adults presented them with four objects, whereas participants only saw two in our design, although it is unclear why this would affect baseline interest. A second important design difference is that in our study, the familiar object was labeled twice before the third critical sentence (e.g., "Look at the book! Ooh, what a nice book! Look at thee *uhh* book/semp!"). This was done following previous research with toddlers (Kidd et al., 2011), to ensure that the novel label would be new to the discourse. An examination of Figure 3 shows that both children and adults' attention was focused on the familiar object prior to the Disfluency Window, likely because that object had just been labeled twice. Adults appeared to maintain greater attention towards this previously-labeled object than children during the disfluency period. It may be that, despite hearing the disfluency, adults nonetheless thought that the speaker would continue on the same topic of discourse.

2.3 General Discussion

We tested monolingual and bilingual children (aged 32 months) and bilingual adults on their predictive use of disfluencies that were either consistent (e.g., "Look at thee *uh* doll!") or inconsistent (e.g., "Look at thee *euh* doll!") with the language of the sentence. Across all groups, listeners increased attention to a novel object rather than a familiar object upon hearing a disfluency. This effect was not modulated by whether the language of the disfluency was consistent vs. inconsistent with the rest of the sentence, or by bilingual participants' language dominance. Results from a control study revealed that bilingual adults could readily discriminate between *uh* and *euh*, showing that the similar response across conditions was not due to an inability to differentiate the English and French disfluencies. Overall, these results suggest that speech disfluencies are a highly robust cue that allow listeners to predict that a speaker is more likely to refer to a novel, discourse-new object.

These studies were designed to test the hypothesis that listeners become sensitive to the predictive nature of disfluencies as they gain associative experience that disfluencies regularly predict the labeling of a novel or discourse-new object. We predicted that hearing typical speech disfluencies in one's ambient language would support listeners to predict that the following word is likely to be novel. Results from our two studies failed to support this hypothesis. Specifically, we predicted that the predictive use of a disfluency would be affected by whether it was language-consistent or language-inconsistent, that bilinguals might be less able to use disfluencies predictively than monolinguals, and that we would see effects of language dominance in bilinguals. However, we did not find evidence that any of these factors affected listeners' ability to use disfluencies predictively.

Our results thus leave open the question of whether and how language experience contributes to listeners' ability to use speech disfluencies predictively. One possibility is that listeners do learn through experience that disfluencies predict novel or discourse-new referents, but that the representation of these disfluencies is inclusive of both native and non-native disfluencies. Under this account, rather than encoding a specific disfluency such as *uh* as being predictive, speakers might group together a range of speech hesitations, for example attending to the lengthening that is characteristic of disfluencies. Even though bilinguals' experience with specific disfluencies is divided across their two languages, they likely encounter a similar total number of disfluencies as monolinguals (or perhaps even more if they encounter more non-native speakers, and produce more disfluencies themselves). Moreover, to the degree that language-consistent and language-inconsistent disfluencies are both perceived as disfluencies, they may be sufficient to activate speakers' knowledge of the association between disfluencies and novel referents.

Another important consideration is that cues correlated with the filled pause might be sufficient to activate such knowledge. In our design, filled pauses were always preceded by a lengthened determiner consistent with the language of the sentence (e.g., *thee uh, thee euh, lee/laa euh, lee/laa uh*). This lengthening was a deliberate choice, given evidence that nearly all filled pauses are preceded by a lengthened determiner (with a vowel alteration in English; Fox Tree & Clark, 1997). Thus, our design cannot rule out the possibility that the lengthened determiner, either by itself or together with the filled pause, signaled that a novel referent would follow, a topic which will need to be investigated in future research. Our results nonetheless demonstrate that the presence of an atypical filled pause (e.g., Language-inconsistent) is not sufficient to disrupt listeners' predictive use of the disfluency.

A second very different possibility is that listeners become sensitive to the predictive nature of disfluencies when they understand them as a marker of a speaker's uncertainty. This framework would also predict that hearing any type of speech disfluency – whether native or non-native – helps listeners understand that a speaker who produces these disfluencies is uncertain, and will likely name a novel object as opposed to a familiar one. Here, children might understand that speech disfluencies signal a speaker's processing difficulty (Clark & Fox Tree, 2002), and infer that a novel referent is what caused the difficulty. Indeed, uncertain speakers tend to be disfluent and make more statements such as "I don't know" (Smith & Clark, 1993). Beginning around age 2, children are increasingly aware that a speaker's verbal cues and actions may indicate a lack of confidence, and this affects their learning from these speakers (Brosseau-Liard & Poulin-Dubois, 2014; Stock, Graham, & Chambers, 2009).

While we did not manipulate our speaker's knowledge state in this study, previous research suggests that speaker knowledge state matters for how listeners interpret disfluencies. For instance, adults' attention to novel objects is not modulated by disfluencies when they are told that a speaker suffers from object agnosia and has difficulty retrieving familiar words (Arnold, Kam, & Tanenhaus, 2007). Relatedly, 3.5 year-old children use disfluencies to predict the labelling of a novel object when hearing a knowledgeable speaker, but not when hearing a forgetful speaker (Orena & White, 2015). Thus, both children and adults appear to modulate their predictive use of disfluencies based on characteristics and knowledge of the speaker.

Sensitivity to a speaker's knowledge state would likely be rooted in children's emerging theory of mind skills. Infants track others' perceptions, beliefs, and goals from early in life (Baillargeon, Scott, & He, 2010; Onishi & Baillargeon, 2005; Yott & Poulin-Dubois, 2016), and these skills become increasingly sophisticated over the preschool years (Astington & Edward, 2010). Previous work indicates that younger children, aged 16–20 months, are not able to use speech disfluencies predictively (Kidd et al., 2011). This could indicate that, between ages 16 and 32 months, there is important development of the social and causal reasoning abilities underlying this ability. It will be important for future research to more explicitly investigate the link between such abilities and the predictive use of disfluencies to more clearly understand how it develops.

Intriguingly for this study, previous research has suggested that bilinguals have accelerated theory of mind development relative to monolinguals (Goetz, 2003; Kovács, 2009). If children's predictive use of disfluencies is rooted in an understanding of a speaker's uncertainty, we might have expected bilinguals to be even more sensitive than monolinguals to the predictive nature of disfluencies, given they may have earlier insight into social signals (Yow & Markman, 2011). However, we found no difference between these two groups. One explanation is that both monolinguals and bilinguals aged 32 months have developed the necessary social reasoning capacity to succeed at our task. It would be important for future studies to test younger bilinguals at an age at which monolinguals previously performed more poorly, around 16–24 months, to determine whether they might show an advantage at this age.

Our study provides an important contribution to the literature on children's predictive use of disfluencies, which had previously been limited to studies of monolingual English-learning children (Kidd et al., 2011; Orena & White, 2015; Owens & Graham, 2016; Owens, Thacker, & Graham, 2017). We replicated these findings with English monolinguals, and extended them to French monolinguals and French-English bilinguals. Moreover, we showed that these groups are flexibly using both Language-consistent and Language-inconsistent disfluencies to make predictions. Our results suggest that children from various language backgrounds are sensitive to language-general characteristics of disfluencies, such as lengthening of the determiner or the filled pause itself. However, an important caveat is that English and French disfluencies (*uh* vs. e*uh*) are somewhat acoustically similar. Filled pauses can sound quite different in other languages: for example, Japanese uses the bisyllabic disfluency *eeto* (Swerts, 1998). An important direction for future research will be to examine whether Japanese learners can use this disfluency predictively from the same age as children in our study, and whether such a disfluency could be used by speakers of English or French, for example. It is possible that speakers of English and French may require the lengthening of a vowel as a key marker and indicator of a hesitation, and thus may not be able to use a bisyllabic disfluency such as *eeto* predictively. Additional cross-language research could better pinpoint what type of knowledge is being acquired across development, and how universal children's sensitivities are.

Importantly, our results show that both toddler and adult listeners, from monolingual and bilingual backgrounds, can flexibly take advantage of speech disfluencies to make predictions about upcoming speech. Moreover, this occurs in real time; listeners increase their attention to a novel object while hearing the disfluency, prior to hearing the noun. Prediction during language comprehension is an important component of efficient language processing, especially for young learners (Lew-Williams & Fernald, 2007; Rabagliati, Gambi, & Pickering, 2016). Since disfluencies occur frequently in human language (Shriberg, 2001), the ability to use them for prediction could make an important contribution to language acquisition. Our evidence shows that children are flexible across different realizations of disfluencies, even those that are not typical of their everyday language input. This suggests that children can interpret disfluencies from interlocutors across a wide variety of backgrounds, including non-native speakers who may produce atypical disfluencies. Adults also showed a general disfluency effect and increased their attention to the novel object upon hearing a disfluency, although they gazed more overall at the familiar objects than children. This reflects a possible attentional strategy favoring known and/or previously-mentioned objects.

Converging with previous research (Arnold et al., 2004; Kidd et al., 2011), our findings confirm that disfluencies direct listeners' attention to novel objects. Our results also showed that, upon hearing the novel label, attention to the novel object was maintained relative to when a familiar label was heard. However, this finding was somewhat less robust for bilingual children, replicating previous findings (Byers-Heinlein & Werker, 2009; 2013; Kalashnikova, Mattock, & Monaghan, 2015). What remains untested is how the presence of disfluencies contributes to subsequent association of the novel object with a novel label. On one hand, disfluencies direct attention to novel objects, which could help children associate the novel objects with an upcoming novel word. On the other hand, children may avoid learning words that follow a disfluency, if they are less willing to learn from speakers who have been unreliable or verbally inaccurate (Brooker & Poulin-Dubois, 2013; Kim, Kalish, & Harris, 2012). Studies investigating other phenomenon have found cases where children can identify the referent of a novel word, but have difficulty forming a word-object association (Horst & Samuelson, 2008). If disfluencies paradoxically increase attention to a novel object while decreasing children's willingness to associate that object with a novel label, this provides additional evidence for a processing of disfluencies that is based on children's understanding of a speaker's knowledge. We are addressing this possibility in an ongoing study with monolingual and bilingual children.

Given the growing literature on bilingual children's processing of speech, the investigation of speech disfluencies is a timely topic that illuminates the mechanisms allowing listeners to comprehend everyday speech. Based on the findings in this study, caregivers and educators should be made aware that some natural imperfections in speech may be beneficial during comprehension. Although our *uh*s and *euh*s are often deemed peripheral to communication, our findings add to evidence that the presence of disfluencies in natural speech is useful for children of different language backgrounds to make predictions about upcoming words during listening. Without hesitation, we contend that speech disfluencies could be an important element of children's language development.

Table 1

Familiar-novel object pairs in English and in French

English Order		French Order		
Familiar	Novel	Det.	Familiar	Novel
book	semp	le	bas [sock]	sèpe
bread	mog	la	bouche [mouth]	taque
cheese	gorp	la	chaise [chair]	moube
cookie	posha	le	chapeau [hat]	brussin
cow	sarl	la	cuillère [spoon]	dimotte
cup	blat	la	dent [tooth]	shème
dog	koob	le	lait [milk]	fide
doll	rel	le	lit [bed]	juve
frog	dak	la	main [hand]	froise
key	mip	la	maison [house]	télue
pig	juff	le	nez [nose]	migue
plate	zel	le	pied [foot]	vuche
stroller	voopi	la	pomme [apple]	kète
toothbrush	glindle	la	porte [door]	tine
truck	shob	le	soulier [shoe]	pafli
window	teeba	le	verre [glass]	triffe

Note. Familiar and novel words in English and French. Masculine (le) and feminine (la) determiners (Det. column) are indicated for familiar-novel pairs in French. English translations of French words are indicated in square brackets. Familiar and novel words in English and French were not translations, to ensure that words were equally familiar or novel to both monolinguals and bilinguals (who saw both English and French orders).

Table 2

Disfluency Window analyses of looking to the novel object on Fluent, Disfluent Language-consistent, and Disfluent Language-inconsistent trials compared to chance (0.5) in Study 1 (monolingual and bilingual children) and Study 2 (bilingual adults)

	Monolinguals					Bilinguals				
	M	*SD*	*t*	*p*	*d*	*M*	*SD*	*t*	*p*	*d*
Study 1										
Fluent	.51	.17	0.210	.84	0.04	.50	.18	0.961	.34	0.24
Dis. Lang.-consist.	.61	.21	2.966	.006	0.52	.66	.23	3.801	.001	0.69
Dis. Lang.-inconsist.	.62	.20	3.474	.002	0.61	.65	.26	3.054	.005	0.56
Study 2										
Fluent						.17	.17	–11.14	.001	1.97
Dis. Lang.-consist.						.30	.26	–4.238	.001	0.69
Dis. Lang.-inconsist.						.30	.26	–4.242	.001	0.56

Note. M = Mean. *SD* = Standard deviation. *t* values are comparisons to chance looking (0.5).

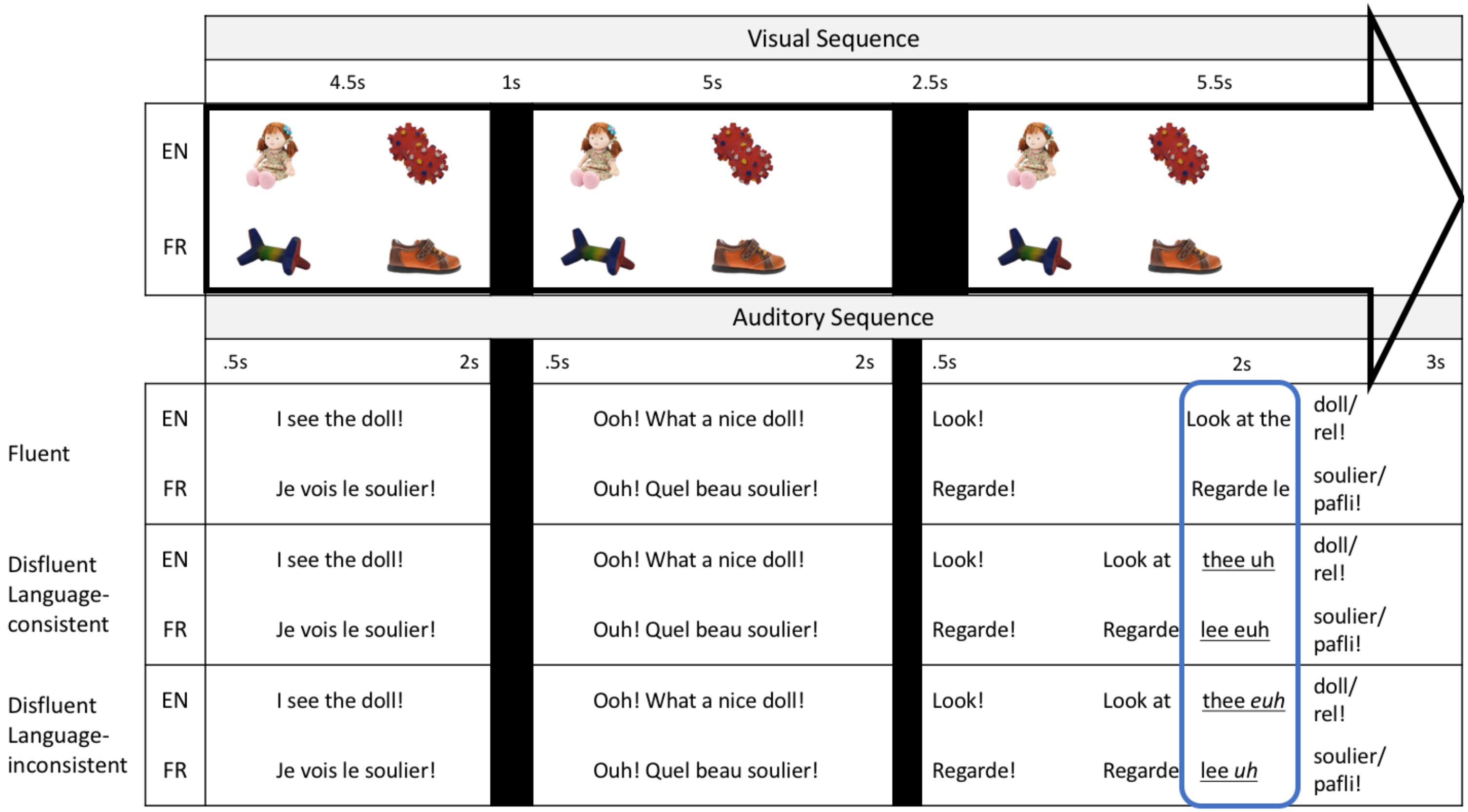


Figure 1. Manuscript 1: The audio-visual structure of trials in English and French, for each Trial type (Fluent, Disfluent Language-consistent, Disfluent Language-inconsistent). EN denotes a sample English trial, whereas FR denotes a sample French trial. The 2 s Disfluency Window of analysis before the onset of the target word in the third sentence is identified by a blue box.

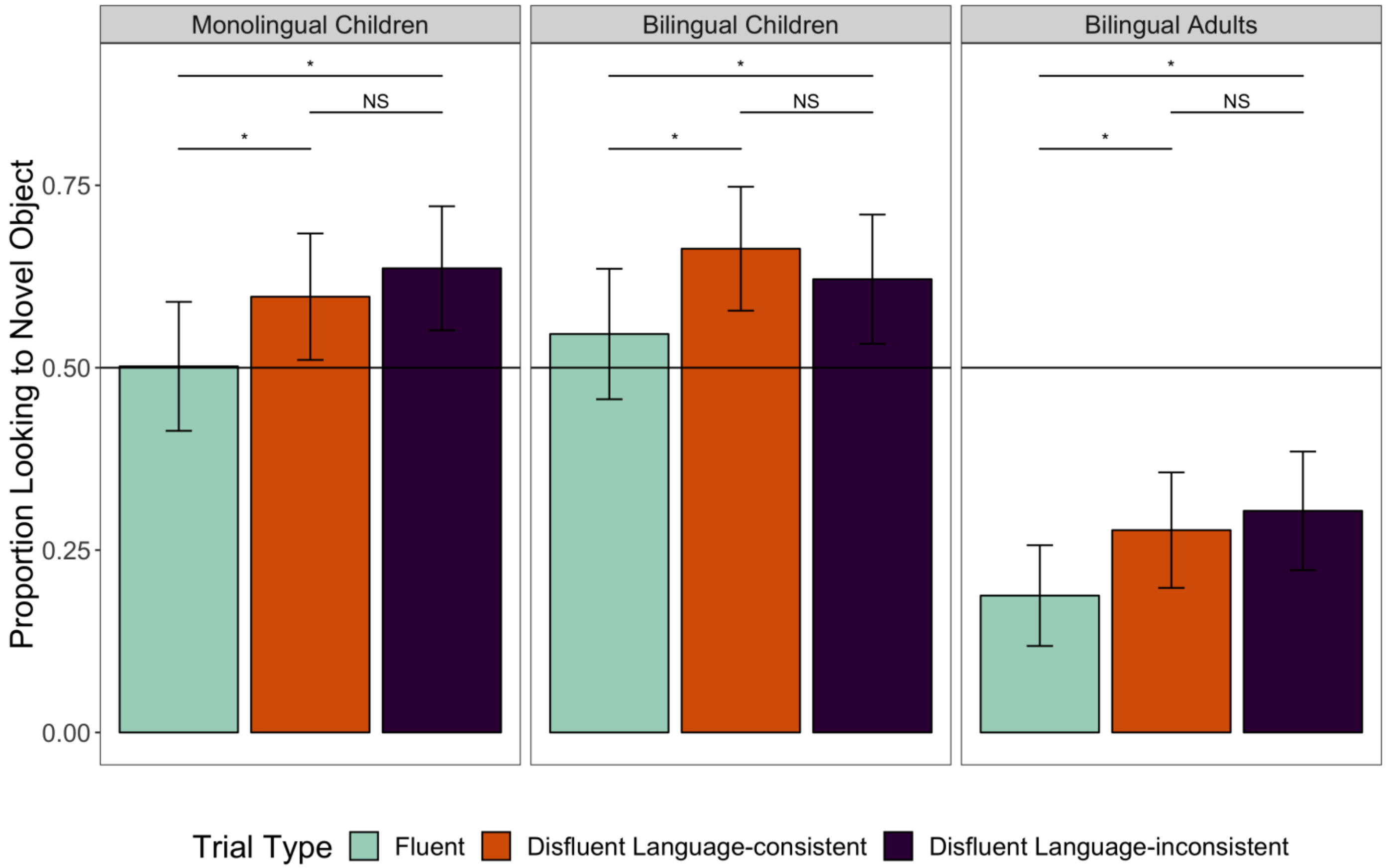


Figure 2. Manuscript 1: Monolingual children, bilingual children, and bilingual adults' looking to the novel object in the disfluency period on Fluent (8 trials), Disfluent Language-consistent (4 trials), and Disfluent Language-inconsistent (4 trials) trials. Error bars represent the standard error of the mean. NS denotes $p > .05$, and asterisks denote $p < .05$.

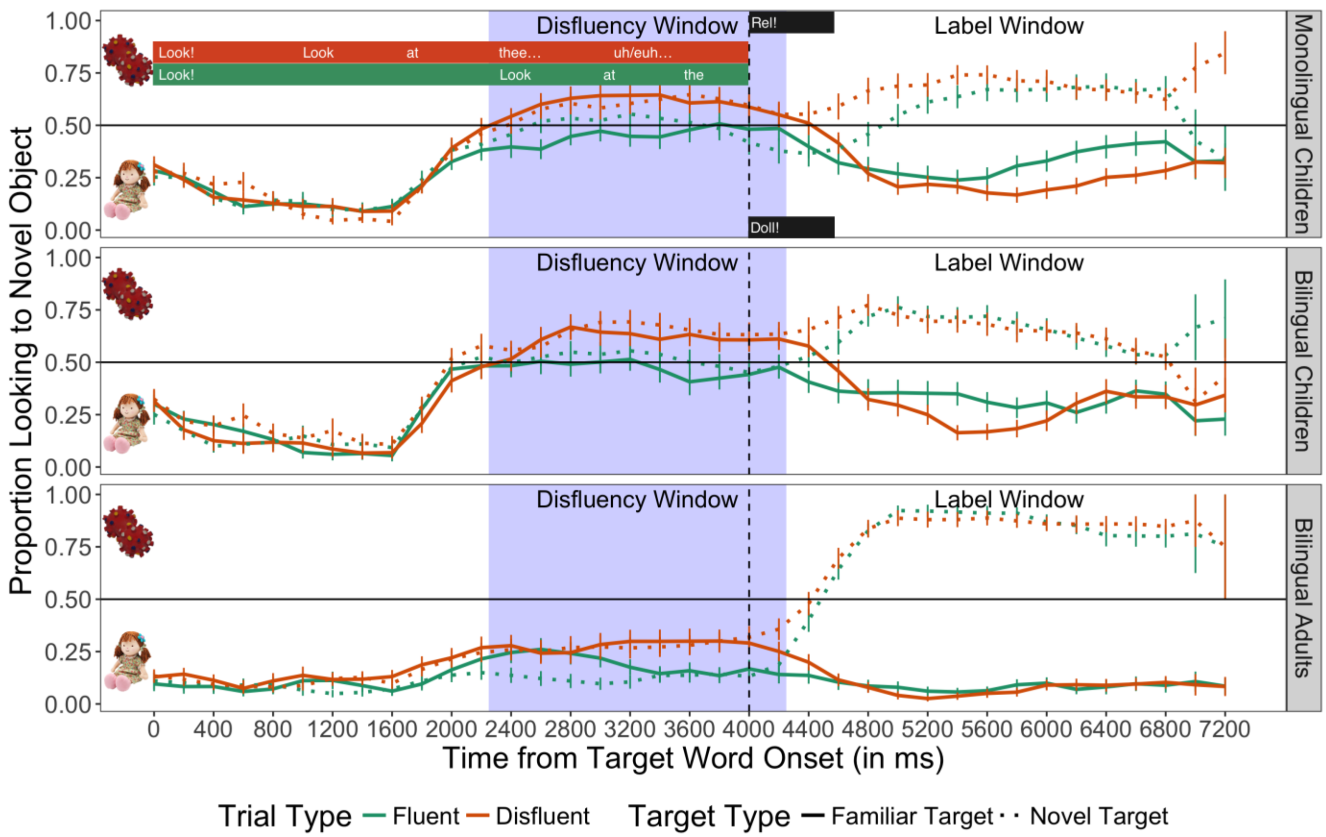


Figure 3. Manuscript 1: Monolingual children, bilingual children, and bilingual adults' looking to familiar and novel target objects following the target word onset. The blue shaded area represents the 2 s Disfluency Window before the target word onset (2250–4250 ms after trial onset), which was shifted by 250 ms in the analysis to account for stimulus response latency. The Label Window corresponds to the area in white following the Disfluency Window (4250–7250ms after trial onset), which was also shifted by 250 ms. The dashed vertical line indicates the target word onset. Chance looking (.50) is represented by the solid horizontal line.

3 CHAPTER 3: Manuscript 2 on Word Learning

Speech is a highly imperfect signal, filled with verbal stumbles such as the filled pauses "uh" and "um" in English (Shriberg, 2001). Hesitation markers such as "uh" and "um" are more generally referred to as *speech disfluencies* (Fox Tree, 1995). Speech disfluencies are commonly encountered by children in the speech of adults. In natural speech, disfluencies often occur before new words, for example speakers tend to hesitate and say "uh" before naming a word that is new to the conversation (e.g., "Look at the, uh... walrus!") (Arnold, Losongco, Wasow, & Ginstrom, 2000). Children likely hear speech disfluencies followed by words that are novel to them on a daily basis. The presence of these disfluencies just before novel words raises the question of whether they hinder or help word learning in young children,

A first possibility is that speech disfluencies hinder novel word learning. Speakers often produce disfluencies prior to a word or topic that is unfamiliar or otherwise difficult to retrieve (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Caramazza, Costa, Miozzo, & Bi, 2001; Jescheniak & Levelt, 1994), and thus children might take disfluencies as a marker of a speaker's uncertainty. Indeed, there is evidence that children understand that speech disfluencies signify a lack of knowledge or uncertainty in a speaker (Orena & White, 2015; Owens & Graham, 2016). Studies that have provided children with both verbal (e.g., stating "I'm not sure") and non-verbal (e.g., shrugging shoulders) signs of uncertainty, unreliability, or lack of confidence have found that children are less willing to imitate (Brosseau-Liard & Poulin-Dubois, 2014; Brosseau-Liard & Poulin-Dubois, 2016) and learn from speakers who have previously been unconfident, unreliable, or verbally inaccurate (Brooker & Poulin-Dubois, 2013; Kim, Kalish, & Harris, 2012). However, to date, no published research has examined whether the presence of disfluencies alone hinder word learning.

A second and perhaps less intuitive possibility is that disfluencies instead support the learning of the label for novel objects. Even though they may be markers of uncertainty, experimental work shows that both children and adults use disfluencies to make predictions about what a speaker will refer to (Arnold et al., 2003; Kidd, White, & Aslin, 2011; Morin-Lessard & Byers-Heinlein, *under review*). For example, in two comprehension studies with children (Kidd et al., 2011; Morin-Lessard & Byers-Heinlein, *under review*), monolinguals and bilinguals saw pairs of one familiar and one novel object on a screen (e.g., a spoon and an oddly-shaped object), with one object labeled either in a fluent (e.g., "Look! Look at the fep!") or disfluent (e.g., "Look! Look at *thee uh… fep*!") sentence. When they heard the disfluency, children increased their attention to the novel object, even before hearing a label. Findings from these studies suggest that children from 28 to 32 months can use hesitations such as "uh" predictively to infer that a speaker will label a novel object as opposed to a familiar one. It is possible that this increased attention to a novel object in the context of disfluencies could boost children in associating that object with a label. Another boost to word learning could arise because disfluencies create a temporal delay in the speech signal (Corley & Hartsuiker, 2011) – a pause that is filled by speech in the case of "uh" – which naturally isolates a novel word to be learned. This could give learners more time to process spontaneous speech, and therefore a better opportunity to encode the following novel word. These aspects of speech disfluencies could facilitate the learning of labels for novel words.

Another possibility is that the effects of speech disfluencies on word learning depend on the learner's background. In a real-time comprehension situation, previous work on children's processing of speech disfluencies has shown that disfluencies affect the timecourse of children's attention to a novel referent, and that such disfluencies are used equally by monolingual and bilingual children (Morin-Lessard & Byers-Heinlein, *under review*). However, in a word learning situation, just after the disfluency, children would also hear a novel word. For monolingual children, this would additionally increase their attention to the novel object through the mutual exclusivity heuristic, which is the assumption that a novel word refers to a novel object rather than a known familiar object. However, bilingual children show less robust use of mutual exclusivity than monolinguals (Byers-Heinlein & Werker, 2009; 2013; Houston-Price, Caloghiris, & Raviglione, 2010), which may be because bilingual children often have two labels for one object compared to monolinguals who only have one label for an object. Thus, in the context of word learning following a speech disfluency, monolinguals have two converging sources of information orienting them towards the novel object (via the disfluency and mutual exclusivity) while bilinguals may need to rely more on the disfluency. The interaction of these different sources of information (speech disfluencies and novel words) with bilingualism has yet to be explored empirically.

Using an eyetracking paradigm, the current research tested whether speech disfluencies would hinder or enhance word learning, and how this might be affected by children's language background. To do so, we investigated 32-month-old children's learning of novel words embedded in fluent versus disfluent sentences. Moreover, we also compared monolingual

(English or French) and bilingual children (learning English or French, and another language). Previous work with children has investigated only how children direct their attention to a novel object compared to a familiar object when hearing a speech disfluency (Kidd et al., 2011, Morin-Lessard & Byers-Heinlein, *under review*). We expanded on this design by first exposing children to novel labels in either fluent or disfluent sentences (similar to previous studies), but adding two test phases in which we measured the degree to which children had formed an association between the novel object and novel label.

Within the learning and test phases, our experimental design comprised several features. Children were presented with two novel words, one which was always in fluent speech and the other which was always in disfluent speech. Trials were presented in a blocked design, with two Learning phases and two Test phases (Learning-Test-Learning-Test). This design was chosen to give sufficient exposure to the novel words to be learned, and to be able to assess the amount of exposure required for novel word learning. Because both a familiar and a novel object appeared on the screen during each trial of the Learning phase, children could use the mutual exclusivity principle (i.e., understanding that objects typically only have one label; Markman & Wachtel, 1988) to infer that a novel label referred to the novel object. On disfluent trials, the disfluency provided an additional cue that a novel word could likely be labeled.

Overall, we predicted that speech disfluencies would facilitate children's learning of novel words due to the processing advantage that they provide over pause-free, fluent speech. However, we also predicted potential differences between monolinguals and bilinguals. Given evidence that monolinguals and bilinguals show similar abilities to use speech disfluencies predictively during comprehension (Morin-Lessard & Byers-Heinlein, *under review*), we expected to replicate this finding during the Learning phase. However, since this task relies on mutual exclusivity, and given that bilinguals are less likely to show mutual exclusivity than monolinguals, (Byers-Heinlein & Werker, 2009; 2013; Houston-Price, Caloghiris, & Raviglione, 2010), we expected that monolinguals would show better overall retention of the novel words in the test phase compared to bilinguals. Thus, we predicted main effects of both Trial Type and Language group during the Test phase, but that there would be no interaction between these two factors. Specifically, we expected that children would learn words better in the presence of disfluent speech compared to fluent speech, and that monolinguals would overall learn words better than bilinguals.

3.1 Method

Participants

Participants were 33 children aged 32 months, who came from monolingual (*n* = 17) or bilingual (*n* = 16) backgrounds. Participants were recruited from a database of eligible participants initially obtained from government birthlists and from daycares, and lived in Montréal, Canada. An additional 17 children were tested but were excluded due to failure to meet pre-established criteria for monolingualism or bilingualism (8), reported low birth weight (2) or other health issues (2), parental interference (2), not providing any looking data on test trials (1), unwillingness to stay in the testing room (1), or equipment error (1). All children included in the final sample were healthy and typically-developing based on parental report. Data collection took place from March to October 2017.

Monolinguals (M_{age} = 32m9d, *SD* = 7d; 14 females) had been exposed to either English (*n* = 13) or French (*n* = 4) from birth, and had no more than 10% exposure to other languages. Bilinguals (M_{age} = 32m13d, *SD* = 13d; 8 females) had been exposed to either English or French since birth and also had regular exposure to an additional language (English, French or another language). Twelve bilinguals had acquired their two languages simultaneously from birth, and four bilinguals had acquired their second language after birth but before age 18 months (M_{AOA} = 13m, range = 12m-18m). The language profiles of bilinguals were: English-French (13), English-Arabic (1), English-Italian (1), and English-Polish (1). The Multilingual Approach of Parent Language Estimate (MAPLE) was used along with an adaptation of the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 2001; Byers-Heinlein et al., *under review*), to evaluate the input from different caregivers across the child's life, and to measure monolinguals' and bilinguals' language exposure (see Appendix B). Bilinguals were exposed to each of their languages at least 25% of the time, and had no more than 10% exposure to other languages. Based on the language of greatest exposure, 10 bilinguals were English-dominant, and 6 bilinguals were French-dominant. No children were dominant in a language other than English or French. On average, bilinguals were exposed to 61.4% of their dominant language (range: 50-75%), and 35.6% of their non-dominant language (range: 28-48%).

Stimuli

Visual Stimuli. Visual stimuli were familiar-novel object pairs (Learning phase) and novel-novel object pairs (Test phase), counterbalanced for side of presentation (left or right), to avoid strategic looking to a specific side for the target (see Kidd et al. (2011) and Morin-Lessard & Byers-Heinlein (under review) for examples of similar familiar-novel object pairings). Images for the two novel objects were selected from the NOUN-2 Database available online (Horst & Hout, 2016), and were judged by two adults to have similar visual salience. Images of familiar objects were obtained from Google Images, and were also selected to have similar visual salience. There were two novel objects, and 8 familiar objects whose labels were expected to be known by children of this age based on vocabulary norms in American English and Québec French (Fenson et al., 2007; Trudeau, Frank, & Poulin-Dubois, 1997). All familiar words in French were of feminine gender. Five attention-grabbing animations were used as attention-getters between each trial to maintain children's interest throughout the study and reorient them to the screen. Object pairs and attention-getters were all presented against a white background.

Auditory Stimuli. Auditory stimuli were recorded by a female native bilingual of Canadian English and Québec French, who spoke in a child-directed manner. Two novel target words were selected, 'moba' and 'voopi'*,* which were chosen because they respect the phonotactic constraints of English and French, match in number of syllables, have distinct onsets and places of articulation, and do not rhyme. Separate recordings were created for the Learning phase and for the Test phase, and both English and French versions were recorded. For the Learning phase, the speaker recorded each word in either a fluent sentence (e.g., "Look! Look at the moba!" / "Regarde! Regarde la moba!") and in a disfluent sentence (e.g., "Look! Look at the *uh* voopi!" / "Regarde! Regarde la *euh* voupie!"). Familiar label stimuli were recorded in the same fluent and disfluent sentence frames. Whether a particular word appeared in fluent versus disfluent sentences was constant within children, and counterbalanced across children. In English sentences the disfluency was the English "uh", and in French sentences the disfluency was the French "euh". Note that these types of disfluent sentences have been used in previous studies in both English (Kidd et al., 2011), and French (Morin-Lessard & Byers-Heinlein, *under review*). For the Test phase, the speaker produced the novel target words in isolation (e.g., "Moba! Moba!" or "Voopi! Voopi!"). Auditory stimuli were normalized in the audio software Praat (Boersma & Weenik, 2012) to 70dB, a comfortable listening level.

The visual stimuli used can be found in Table 3, and additional stimuli details can be found in Figure 4.

Procedure

Children were tested in a sound-attenuated room while sitting on their parent's lap, about 60cm away from a Tobii T60XL eyetracker screen. An experimenter controlled the study from an adjacent room using Tobii Studio Software. To avoid influencing the children's behaviour, all parents were instructed not to interact with their child during the study, wore darkened sunglasses, and listened to music via headphones. The study began following a 5-point experimenter-controlled calibration routine.

Children saw a total of 24 trials, divided into two blocks of trials, which each consisted of an eight-trial Learning phase and a four-trial Test phase. Presenting Learning and Test phases in two blocks allowed for the investigation of learning at two different time points. Each child saw trials either in English or in French, whichever language was native (monolinguals) or dominant (bilinguals). During each Learning phase, children saw pairs of objects that always included one familiar and one novel object, and heard a label for one of the objects. On two fluent trials per block, children heard one of the novel objects (e.g., voopi) labeled in a fluent sentence. On two disfluent trials per block, they heard the other novel object (e.g., moba) labeled in a disfluent sentence. Familiar objects were also labeled on two fluent and two disfluent trials per block, which meant that familiar and novel objects were labeled an equal number of time in the study. The object that was labeled fluently or disfluently was consistent within, but counterbalanced across children. Because of the presence of the familiar object, children needed to infer which object the novel label referred to on these trials.

On each Learning Phase trial, the onset of the auditory stimulus (i.e., the first "Look!") occurred 500ms into each trial, and the novel word onset occurred at 4500ms into each trial. The total duration of each Learning phase trial was 7.5s. In the Test Phase, on 4 test trials in a fixed order (each object appeared twice on the left, twice on the right, counterbalanced across children), children saw the two novel objects side-by-side, and heard a label naming one of them. Each word was labeled on two trials, and the label was repeated twice on each trial (e.g., "Moba! Moba!" and "Voopi! Voopi!"), following a 3000ms silent baseline on each trial, and again 2000ms after the onset of the first label. Trials of the same type (i.e., fluent, disfluent) appeared no more than twice in a row. The total duration of each Test phase trial was 8s. As both novel objects were equally familiar, and labels were presented in isolation (rather than in a disfluent or fluent sentence) children could only succeed in the Test phase – correctly identify the target word – if they had associated the novel label with the novel object during the Learning phase. Each child saw the trials in one of eight study orders (4 English orders, 4 French orders), which counterbalanced the side of presentation of the objects, and determined which of the two novel words was presented in the fluent versus disfluent sentence. Because our main interested was in the effect of Trial type (fluent, disfluent) and Language background (monolingual, bilingual), and because the aforementioned variables already required 8 study orders to counterbalance, each novel object was always paired with the same object. The study lasted 5 minutes in total. The auditory timeline of Learning Phase and Test Phase trials is illustrated in Figure 5.

Parents filled demographic and language background questionnaires either before or after the main study (see Appendix C), and children received an honorary diploma and a gift for partaking in the study.

3.2 Results

The main analysis of interest was the Test phase analysis, which assessed children's learning of the novel words. The secondary analysis of interest was the Learning phase analysis, which allowed us to investigate children's looking strategy during learning. Test phase analyses are reported first, followed by Learning phase analyses.

Looking data towards each object collected were summed across the relevant time window for each trial, and for each area of interest using eyetrackingR for R (Dink & Ferguson, 2015; R Core Team, 2017).

Test Phase Analyses

The dependent variable in the Test phase was the proportion of looking to the target object, which was obtained by dividing the total fixation duration at the target object by the sum of the total fixation duration at both the target and distractor objects. Finally, data for each trial type were averaged for each block. Here, we report the results of pre-planned analyses which focused on comparisons across trial types and comparison to chance (.5) given our counterbalanced design. However, in a later analysis, we also report follow-up analyses that correct for different baseline levels of object salience.

Data were gathered in two blocks of trials, resulting in two blocks of test trials with four trials (2 fluent, 2 disfluent) in each block. We initially planned to include block as a within-subjects variable within the same ANOVA, which would necessitate that participants contribute data from both trial types on both blocks. However, while 32 participants (16 monolinguals, 16 bilinguals) contributed data from both trial types in Block 1, only 26 participants (15 monolinguals, 11 bilinguals) contributed data from both trial types in Block 2, suggesting that many children had lost interest in the task by the second block. Thus, we report results from Block 1 as our primary analysis. For completeness, we have also reported results from Block 2, although we note that they should be interpreted with caution.

For both blocks, the window of analysis spanned from 250ms after the target word onset to 4000ms later, and thus started at 3250ms and ended at 7250ms after the beginning of the trial. The window of analysis was shifted by 250ms as in previous studies (Kidd et al., 2011; Morin-Lessard & Byers-Heinlein, *under review*) to account for the time it takes to launch an eye movement. This 4s window allowed us to assess learning of novel words following one (Block 1) or two (Block 2) learning phases.

Block 1. Test phase results for Block 1 are displayed in Figure 6. To investigate whether there was an effect of disfluent speech and language background on children's learning of words, a 2 (Trial type: Fluent, Disfluent) x 2 (Language group: monolingual, bilingual) mixed ANOVA was performed. There was no significant main effect of Trial type, $F(1, 30) = 2.16$, $p = .15$, $\eta^2 = .003$, no main effect of Language group $F(1, 30) = 0.09$, $p = .77$, $\eta^2 = .067$, and no interaction between Trial type and Language group, $F(1, 30) = 0.00$, $p = .99$, $\eta^2 < .000$. Follow-up single-samples *t*-tests comparing attention to the labeled object to chance (.5) did not reveal any evidence that children in either group (monolinguals, bilinguals) recognized the words during the Test phase, whether they had been taught in fluent or disfluent sentence frames (*p*s > .05. See Table 4 for results). Unsurprisingly, neither group showed a significant difference in looking for the Fluent vs. the Disfluent trials (*p*s > .05). While analyses comparing trial types did not reach statistical significance, we note that Block 1 means for monolinguals and bilinguals are in the direction of more looking to the target on disfluent trials compared to on fluent trials.

Block 2. Test phase results for Block 2 are displayed in Figure 6. As in Block 1, a 2 (Trial type: Fluent, Disfluent) x 2 (Language group: monolingual, bilingual) mixed ANOVA was performed to investigate the effect of disfluent speech and language background on word learning, with looking after hearing the target word as the dependent variable. Again, there was no significant main effect of Trial type, $F(1, 24) = 0.16$, $p = .69$, $\eta^2 = .007$, no main effect of Language group, $F(1, 24) = 0.14$, $p = .72$, $\eta^2 = .006$, and no interaction between Trial type and Language group, $F(1, 24) = 2.46$, $p = .13$, $\eta^2 = .093$. Comparisons to chance again found no evidence that children in either group recognized either of the novel words at test (*p*s > .05. See Table 4 for results). Bilinguals showed no difference in their looking between Fluent trials (M = .55, SD = .39) and Disfluent trials (M = .44, SD = .26), $t(10) = 0.671$, $p = .52$, $d = 0.20$. Monolinguals looked less on Fluent trials (M = .44, SD = .32) than on Disfluent trials (M = .62, SD = .27), $t(14) = -1.876$, $p = .08$, $d = -0.48$, a difference which was marginal but not statistically significant. Although there is no strong statistical evidence to conclude that children learned either word, this might be some evidence that monolinguals performed better on disfluent trials compared to fluent trials in Block 2.

Salience-corrected analyses. To further investigate the weak evidence of word learning in the Test phase, we investigated looks to each novel object separately during the silent baseline preceding target labeling, which spanned from the beginning of the trial at 0ms, to 3000ms later (see Figure 5). An average proportion of looking to the moba compared to the voopi was calculated for each block. In Block 1, the average proportion of looking to the moba was .61 (SD = .10), which was significantly above chance (.50), $t(32) = 6.427$, $p < .001$. In Block 2, the average proportion of looking to the moba was .65 (SD = .18), which was also significantly above chance, $t(31) = 4.825$, $p < .001$. This suggested that, independent of labeling, children had a baseline preference to look at the moba.

To account for the preference to look at the moba in the Test phase, we calculated a target looking difference score for each participant by subtracting their preference looking at baseline (from 0ms-3000ms) from the proportion of looking to the target after it was labeled (from 3250ms-7250ms). This difference score allowed us to assess whether looking to a given target increased or decreased following target labelling. Specifically, a positive difference score reflected an increase in looking to the target object following its labelling, whereas a negative difference score reflected a decrease in looking to the target following its labelling, relative to children's looking during silence. Since there was an unequal number of participants contributing trials in both blocks, results from Block 1 and Block 2 were analyzed separately.

Similarly to Test phase analyses before baseline looking corrections, we initially planned to include block and trial type as within-subjects variables within the same ANOVA, but participants did not contribute data from both trial types in both blocks. In Block 1, we were able to run an ANOVA with trial type and language group, as an equal number of participants from each language group (16 monolinguals, 16 bilinguals) contributed data from both trial types.

However, by Block 2, participants began to lose interest in the task: In Block 2 on fluent trials, 16 monolinguals contributed data compared to 13 bilinguals. In Block 2 on disfluent trials, 15 monolinguals contributed data compared to 11 bilinguals. Due to children's decreased interest in Block 2, we report results from Block 1 as our primary analysis. For completeness, we have also reported *t*-test results from Block 2, but we note that they should be interpreted with caution.

Difference score results for both Block 1 and Block 2 are displayed in Figure 7.

Block 1. To investigate whether there was an effect of disfluent speech and language background on children's baseline-corrected looking to the target, a 2 (Trial type: Fluent, Disfluent) x 2 (Language group: monolingual, bilingual) mixed ANOVA was performed. There was no significant main effect of Trial type, $F(1, 30) = 0.62$, $p = .44$, $\eta^2 = .006$, no main effect of Language group $F(1, 30) = 1.03$, $p = .32$, $\eta^2 = .024$, and no interaction between Trial type and Language group, $F(1, 30) = 0.09$, $p = .76$, $\eta^2 < .001$. Follow-up single-samples *t*-tests comparing attention to the labeled object to chance (0) did not reveal evidence that children in either group (monolinguals, bilinguals) recognized the words during Block 1 of the Test phase, either when taught in fluent or disfluent sentence frames (*p*s > .05; see Table 4 for results after baseline looking corrections). While analyses compared to chance did not reach statistical significance, we note that the Block 1 looking mean on fluent trials for bilinguals was positive and marginally significant. These results indicate that children did not show comprehension of the new words in Block 1, with the exceptions of bilinguals, who may have looked more to the target object on fluent trials.

By-items analyses. Given the lack of strong evidence of comprehension at test even following baseline corrections, we also investigated baseline-corrected looking to the target object, this time for each individual novel object (moba vs. voopi). There was no effect of trial type or language group, therefore data were collapsed over these factors, *p*s > .05. In Block 1, the average target looking difference score to the moba was negative ($M = -.06$, $SD = .28$) and did not significantly differ from chance (0), $t(31) = -1.32$, $p = .20$. In contrast, the average target looking difference score to the voopi was positive ($M = .19$, $SD = .28$) and was significantly above chance (0), $t(31) = 3.92$, $p < .001$. Given that children already had a strong preference to look at the moba before it was labeled, it is unclear whether children were not able to learn its label, or did learn it but could not demonstrate their learning due to ceiling effects. In comparison, children's looking to the voopi object increased after it was labeled compared to before it was labeled, suggesting that they possibly showed comprehension of the new word in Block 1 of the Test phase.

Block 2. One-sample *t*-tests were used to compare attention to the labeled object to chance (0) in Block 2. These results did not reveal evidence that children recognized words in Block 2 of the Test phase, whether they were taught in fluent or disfluent sentence frames (*p*s > .05; see Table 4 for results after baseline looking corrections). We note, however, that the Block 2 looking mean on fluent trials for bilinguals was negative and marginally significant, suggesting that bilinguals may have looked less to the target object on Fluent trials, a finding that is in the opposite direction as the one found in Block 1. We highlight that these results should be interpreted with caution due to lower levels of participant interest in Block 2, and resulting missing data.

By-items analyses. As in Block 1, since there was no effect of trial type or language group (*p*s > .05), data were collapsed over these factors, and we investigated baseline-corrected looking to the target object by item (moba vs. voopi). Analyses revealed that the average target looking difference score to the moba in Block 2 was negative (M = -.09, SD = .31) and did not significantly differ from chance (0), $t(26) = -1.57$, $p = .13$. In comparison, the average target looking difference score to the voopi was positive (M = .07, SD = .31), but this difference score also did not significantly differ from chance, $t(26) = 1.16$, $p = .26$. As in Block 1, analyses in Block 2 highlight greater looking to the moba at baseline than following its labelling, and greater looking to the voopi following its labelling than at baseline. However, difference scores for both objects in Block 2 did not differ from chance, making it unclear whether children showed comprehension of the new words in the Test phase.

Learning Phase Analyses

Given children's failure to show clear comprehension of the new words in the Test phase, it was important to examine children's looking to the novel object versus the familiar object in the Learning phase. There were two time periods of interest. First, we were interested in their looking during the Disfluency window, which occurred 2 seconds *prior* to the naming of a target object, specifically between 2750ms–4750ms allowing for 250ms for the launching of eye movement. The dependent variable in the Disfluency window was the proportion of looking to the novel object (i.e., looking novel object/(looking novel object + looking familiar object)). Looking to the novel object during this window would replicate previous findings (Kidd, White,

& Aslin, 2011; Morin-Lessard & Byers-Heinlein, *under review*). Second, we were interested in children's looking during the Label window, which occurred *after* the naming of the target object and lasted 2 seconds from 4750-6750ms, again allowing 250ms for children's eye movement. The dependent variable in the Label window was the proportion of looking to the target object (i.e., looking target object/(looking target object + looking distractor object)). This allowed us to determine whether children were attending to the target object at the moment of hearing its label. Data for each trial type for both the Disfluency and Label windows were averaged for each block.

As in the Test Phase, because there was an unequal number of participants contributing trials in both blocks, we also analyzed results from Block 1 and Block 2 separately. Children appeared to be more attentive during the Learning phase than in the Test phase, as all participants contributed data from both trial types in Block 1, and all but one did so in Block 2.

As in the Test phase, looking data for the Learning phase was obtained for looks within areas of interest spanning 2cm around each object, which were a novel and a familiar object.

Block 1.

Disfluency window. See Figure 8 for Block 1 results. A 2 (Trial type: Fluent, Disfluent) x 2 (Language group: monolingual, bilingual) mixed ANOVA was performed to investigate the proportion of looking to the novel object during the disfluency period. The ANOVA revealed a marginally significant effect of Trial type $F(1, 31) = 3.17$, $p = .08$, $\eta^2 = .093$, but no main effect of Language group, $F(1, 31) = 0.25$, $p = .62$, $\eta^2 = .008$. There was no interaction between Trial type and Language group, $F(1, 31) = 1.43$, $p = .24$, $\eta^2 = .044$. Follow-up paired-samples *t*-tests with monolinguals revealed no difference in their looking between Fluent trials ($M = .51$, $SD = .14$) and Disfluent trials ($M = .49$, $SD = .12$), $t(16) = 0.428$, $p = .67$, $d = 0.10$. However, the same analysis revealed that bilinguals looked at the novel object marginally significantly more on Fluent trials ($M = .52$, $SD = .19$) than on Disfluent trials ($M = .41$, $SD = .14$), $t(15) = 1.916$, $p = .07$, $d = 0.48$. For monolinguals, neither looking to the novel object on Fluent trials nor on Disfluent trials differed from chance, all *p*s > .05. For bilinguals, looking to the novel object on Fluent trials did not differ from chance, however looking to the novel object on Disfluent trials was significantly below chance. See Table 5 for results. These analyses suggest that monolinguals did not look more at the novel object upon hearing a disfluency, and that bilinguals looked more at the familiar object than the novel object on disfluent trials. Thus, unlike in previous work, we did not find evidence in this paradigm that the disfluency itself boosted attention to the novel object.

Label window. Results for Block 1 are displayed in Figure 9. A 2 (Trial type: Fluent, Disfluent) x 2 (Target type: Familiar, Novel) mixed ANOVA revealed no main effect of Trial type, $F(1, 31) = 0.17$, $p = .68$, $\eta^2 = .002$, no main effect of Language group, $F(1, 31) = 1.27$, $p = .27$, $\eta^2 = .064$, and no interaction between Trial type and Language group, $F(1, 31) = 1.58$, $p = .22$, $\eta^2 = .060$. Follow-up single-sample *t*-tests showed that monolinguals looked at the novel target significantly above chance on both fluent and disfluent trials, showing that they identified the target referent. Bilinguals looked significantly above chance on Disfluent trials, but not on Fluent trials. However, neither group showed a significant difference between the two trial types, all *p*s > .05. See Table 6 for detailed results.

Block 2.

Disfluency window. See Figure 8 for Block 2 results. Similar to Block 1, a 2 (Trial type: Fluent, Disfluent) x 2 (Language group: monolingual, bilingual) mixed ANOVA revealed no main effect of Trial type, $F(1, 30) = 0.41$, $p = .53$, $\eta^2 = .013$, no main effect of Language group $F(1, 30) = 2.47$, $p = .13$, $\eta^2 = .076$, and no interaction between Trial type and Language group, $F(1, 30) = 0.19$, $p = .67$, $\eta^2 = .006$. Surprisingly, for both groups, looking to the labeled novel target was significantly *below* chance for both Fluent and Disfluent trials, in that they were looking more to the familiar object than the novel object upon hearing the disfluency. Moreover, there was no significant difference in looking between the two trial types. See Table 5 for detailed results. Together, these results suggest that monolinguals and bilinguals looked more to the familiar object compared to the novel object when hearing a speech disfluency. Thus, there was no evidence for a disfluency effect, which would be characterized by more looking to the novel object specifically on disfluent trials.

Label window. Results for Block 1 are displayed in Figure 9. Once again, a 2 (Trial type: Fluent, Disfluent) x 2 (Target type: Familiar, Novel) mixed ANOVA revealed no main effect of Trial type, $F(1, 30) = 0.63$, $p = .43$, $\eta^2 = .020$, no main effect of Language group, $F(1, 30) = 1.62$, $p = .21$, $\eta^2 = .051$, and no interaction between Trial type and Language group, $F(1, 30) = 0.10$, $p = .75$, $\eta^2 = .003$. Here, both groups looked to the novel target statistically significantly above chance, with the exception of bilinguals on the Disfluent trials whose looking was marginally significant. Neither group showed a significant difference in looking between trial types, all *p*s >

.05. These results indicate that, in general, children successfully looked at the target word when it was labelled in Block 2, regardless of whether it was in a fluent or disfluent sentence frame. See Table 6 for detailed results.

3.3 Discussion

The goal of this research was to investigate whether the presence of speech disfluencies just before novel words hinders or helps novel word learning in 32-month-old children of monolingual and bilingual backgrounds. If disfluencies hinder novel word learning, we predicted that this might be due to children being reluctant to learn from situations in which a speaker is uncertain, such as when a speaker produces disfluencies. In this case, we predicted that children would learn the word in the fluent but not in the disfluent condition. On the other hand, if disfluencies facilitate novel word learning, this may be due to the increased attention to the novel word as children hear a disfluency. In this case, we predicted that children would learn the word in the disfluent but perhaps not in the fluent condition.

Our results did not match either set of predictions. Unexpectedly, children – both monolinguals and bilinguals – showed little evidence of learning the word in *either* condition. In other words, children failed to learn the new words whether they were presented in fluent or disfluent sentences. This contrasts with one unpublished study that I am aware of based on a personal communication, which used live interactions with 3- and 4-year-old children, and investigated children's learning of novel words preceded or not by disfluencies (White, Nilsen, & Riemersma, *under review*). Children were presented with pairs of novel and familiar objects in a live training session with an experimenter, who introduced the novel objects to children. Some novel objects were preceded by a disfluency and others not. Children were then asked to select the novel object. Novel word learning was then assessed via preferential looking, just as in the current study. This study also found no difference between the two conditions, but in contrast to our results, both 3- and 4-year-olds learned the words presented in both the fluent and the disfluent conditions. Although there are obvious differences between our computer-based design and this previous live interaction design, the current findings suggest that the presence of disfluencies could be problematic for word learning before age 3, and that word learning in the context of disfluent speech improves later in development.

While children did not show clear evidence of learning the novel words in this study, planned analyses for the Test phase revealed a possible trend towards more looking to the target on disfluent trials in both Block 1 and Block 2 for monolingual children, and in Block 1 for bilingual children. Further analyses suggested that children had a baseline preference for one of the two objects (the moba). Analyses that corrected for baseline preferences showed that bilinguals were marginally above chance in Block 1 for fluent trials only, but marginally below chance for fluent trials in Block 2. By-object analyses showed some evidence for learning the "voopi" label but not the "moba" label, but it was not clear whether this was affected by fluency condition, and ceiling effects may have masked learning of the "moba" label. Overall, we do see some trends in the data, but in general, these do not reach conventional levels of statistical significance (.05), are inconsistent in their direction, appear to be strongly influenced by object salience, and at times have low power due to infants losing interest in the Test phase of Block 2. Thus, we do not find compelling evidence that infants learned words in this study.

The important question in light of these findings is: why would children fail to learn the novel words? Novel word learning has been reported in monolinguals our participants' age and younger (see book chapter by Werker & Fennell, 2004) and established in looking tasks as early as 12-14 months (e.g, MacKenzie, Curtin, & Graham, 2012; Werker, Cohen, Lloyd, Casasola, & Stager, 1998), with some studies also showing word learning via disambiguation (Bion, Borovsky, & Fernald, 2013) and via associative word learning (Byers-Heinlein, Fennell, & Werker, 2013) in bilinguals younger than the monolinguals tested in this study. Children in our study were aged 32 months and were exposed to a total of 16 learning trials on which novel words were labeled on 8 trials (4 trials for each novel word), which should have been sufficient for children to learn the novel words. Yet, there was no evidence from either Block 1 or Block 2 that children learned either of the two novel words. What was noticeably different between the two blocks is that children's attention was greater in Block 1 than in Block 2, suggesting that their attention decreased over the course of the study. This suggests that children were learning something about the task itself which led them to become disinterested in the study, although we did not find evidence that they learned the novel words that this study was designed to teach them.

To further explore other factors that might explain our findings, we examined children's behaviour during the Learning phase. A first important question was whether children were ultimately able to direct their attention to the target object on learning trials. Overall, we found that children were generally successful at identifying the target referent prior to the end of the trial. The exception was bilinguals' looking to the target object on fluent trials in Block 1, which did not significantly differ from chance looking. Thus, it appears that children were able to identify the correct referent during the Learning phase, but did not retain the word-object mapping in the Test phase. Recent theories of word learning have emphasized referent selection and word retention as two distinct parts of the word learning process (Bion, Borovsky, & Fernald, 2013; Horst & Samuelson, 2008; McMurray, Horst, & Samuelson, 2012; Samuelson & McMurray, 2017). For fluent trials in the current study, children would have needed to rely on heuristics such as mutual exclusivity to orient to the correct novel object (Markman & Wachtel, 1988). It appears that children in our study were able to use this heuristic, although bilinguals in Block 1 may not have been able to use this heuristic as reliably as monolinguals. Nonetheless, despite looking at the correct object during training, children did not appear to have retained the word-object link at test.

One other source of information on disfluent trials during the Learning phase was the disfluency itself, which we expected would draw children's attention to the novel object on Disfluent trials. However, while this phenomenon has been replicated with monolinguals and bilinguals of this age (Kidd, White, & Aslin, 2011; Morin-Lessard & Byers-Heinlein, *under review*), we did not replicate this effect in the current study. That is, children did not increase their attention to the novel object upon hearing the disfluency. Moreover, in Block 2, both monolinguals and bilinguals looked significantly more at the familiar object than the novel object. This was observed both for fluent and disfluent trials, indicating that children were overall more interested in the familiar object during the second half of the study. It is unclear why we did not find a disfluency effect showing more looking to the novel object as listeners heard a disfluency. There are several possible explanations for our failure to replicate this effect. The increased looking to the familiar object for both language groups in Block 2 might simply reveal preferential looking to objects that are familiar over novel, which can occur when exposure to a familiar object is brief (Hunter, Ross, & Ames, 1982). In this study, familiar objects (e.g., banana) only appeared once in each block, whereas each of the two novel objects appeared four times in each block(e.g., voopi). Indeed, we note that our study was designed as a word learning study, which necessitated a somewhat different experimental design from previous studies focusing on real-time word comprehension. One difference is that we used single-utterance trials, which were either fluent (e.g., "Look! Look at the moba!") or disfluent (e.g., "Look! Look at the *uh* moba!"). In previous studies (Kidd, White, & Aslin, 2011; Morin-Lessard & Byers-Heinlein, *under review*), the familiar object was always labeled twice in two separate utterances (e.g., "Look at the book!", "Ooh, what a nice book!") before the critical sentence that we used in the present study (e.g., "Look! Look at the moba!" or "Look! Look at the *uh* moba!"). These two prior utterances established novel words not only as novel in relation to familiar words, but also as new to the discourse. The decision to only use the critical sentence in the present study was to simplify our design and make sure that the study would not be too long. However, since previous studies found that an important driver of the disfluency effect is discourse novelty (as opposed to solely object novelty; Owens & Graham, 2016), this modification to our design may have prevented us from replicating the disfluency effect found in prior studies with children. It may also explain why we found more looking to the familiar object in children in Block 2. At the same time, this possibility does not explain why children also failed to learn the word in the fluent condition.

In sum, on both fluent and disfluent trials, there was an apparent gap between children's referent selection and word retention – children gazed at the novel object once it was labeled, but did not appear to retain the word-object link. In our study design, there are several potential explanations for why children failed at the retention step of word learning.

First, whereas the Learning phase presented children with one novel and one familiar object, the Test phase presented children with the two novel objects. Recognizing the novel words in this situation may have been too cognitively demanding for children due to their fragile knowledge of those words (Kucker, 2014). Mapping the novel words to the novel objects may have been particularly difficult due to weak links between the two, which may require more exposure over the course of development (Kucker, McMurray, & Samuelson, 2015), beyond the testing session in the lab. It is possible that children did make some preliminary word-object association, but were unable to show this knowledge in the Test phase.

Second, the salience of the objects in our study – both familiar and novel objects – may also have affected children's ability to learn the novel words. For familiar objects, although the salience of objects typically concerns how objects look in relation to their surroundings (e.g., their contrast and luminance), objects can also be more salient by virtue of attracting more attention due to other characteristics (e.g., highly familiar objects). This object salience is thought to make it difficult for children to disengage their attention from those objects. In a recent study on the effect of familiar object salience on the learning of novel words (Pomper & Saffran, 2018), 3-year-olds learned novel words above chance when they had been presented next to familiar objects low in salience (e.g., dull-coloured household objects) during learning, but not when they had been presented next to familiar objects high in salience (e.g., brightly-coloured animals, vehicles, food). Although this study presented images that were similar in quality and contrast, familiar objects were selected not based on salience but because they were highly familiar to children of our study's age (Fenson et al., 1993; Trudeau, Frank, & Poulin-Dubois, 1997). The eight familiar objects used in this study were all used in Morin-Lessard & Byers-Heinlein (*under review*; i.e. Manuscript 1) with the exception of *banana* since bilinguals were tested in each of their languages, and the French translation *banane* is too similar). However, there were 16 familiar objects in each order in Morin-Lessard & Byers-Heinlein, (*under review*), as a novel word was presented on each trial and was not presented repeatedly for the purposes of learning, as in the present study.

For novel objects, pairing a more salient novel object with a less salient one may also have been problematic in a word learning situation. In the present study, the extent to which children preferred one object (the moba) over the other (the voopi) was surprising, since the two novel objects had been judged by two adults to have similar salience. Moreover, the 32-month-olds in our study were much older than theories of when object salience might override other factors in learning (Pruden et al., 2006). Our results reflect how object salience can disproportionately affect gaze during word learning tasks, and are in line with theories positing that object salience can influence children's linking of words with objects (Hollich et al., 2000). Together, our findings also highlight that novel object pairs destined for word learning studies should be piloted by children – as opposed to adults – for salience, and should always be considered in follow-up analyses. Finally, a careful consideration of how object salience may interact with novel word learning in the context of disfluent speech will be important for future research.

A third explanation is that the "newness" quality of novel objects may have decreased over the course of the study. Our design presented children with the same two novel objects repeatedly throughout the study to maximize learning. However, the visual novelty of the novel objects may have diminished as children viewed more trials (Mather & Plunkett, 2009), and the salience of the novel object against the familiar object may have decreased (Kucker, McMurray, & Samuelson, 2015). The repetitive nature of the word learning design combined with the diminishing relative novelty of novel words may also explain why children's attention has decreased from the first to the second block. This is especially true for the more challenging Test trials which were used to assess word learning, and always were presented after the Learning trials.

Finally, and in line with previous work on speaker unreliability (Brooker & Poulin-Dubois, 2013; Kim, Kalish, & Harris, 2012), children's novel word learning difficulty may have been amplified by the presence of speech disfluencies. Indeed, if children understand that a speaker is disfluent because he or she is unsure about the label of a word, but hear the same speaker repeatedly associate the same novel word with the same novel label, they may not see the value in using speech disfluencies to predict that a novel object is likely to be labeled. The speaker's repeated disfluencies prior to the same novel word may have diminished their confidence in the speaker and increased their perceived unreliability in the speaker, and have prevented their learning of that word. Children's lack of confidence in the speaker from disfluent sentences may have had a carry-over effect on fluent trials, thereby hindering children's overall novel word learning. We are currently testing this possibility by testing children in a follow-up fluent-only condition, to examine whether word learning improves in this paradigm in the complete absence of disfluencies.

The question of whether disfluencies play a role in word learning – whether detrimental or beneficial – is consequential. Disfluencies are produced frequently in natural speech, which makes it crucial to understand how they may impact children's word learning. The goal of the current study was to understand whether speech disfluencies would help or hinder word learning relative to fluent speech. We found that in our paradigm, children showed an ability to locate the appropriate referent during both fluent and disfluent learning trials, although we found no evidence that the disfluency itself boosted attention to the novel object. Moreover, children did not appear to retain a mapping between word and object at test in either the fluent or disfluent condition. While these results cannot specifically address the role of disfluencies in word learning, they do point to the fragility of word learning even at age 32 months. Future research that includes additional controls, such as between-subjects manipulations of fluency, may help to better understand how speech disfluencies impact word learning.

Table 3

Familiar-novel object pairs in the Learning phase in English and French

Det.	Familiar	Novel
	English/French	
the/la	door/porte	moba/moba
the/la	frog/grenouille	moba/moba
the/la	key/clé	moba/moba
the/la	spoon/cuillère	moba/moba
the/la	apple/pomme	voopi/voupie
the/la	banana/banane	voopi/voupie
the/la	chair/chaise	voopi/voupie
the/la	mouth/bouche	voopi/voupie

Note. Familiar and novel words in English and French. Determiners (la) and nouns in French were all feminine.

Table 4

Test phase analyses of looking to the target before and after baseline looking correction on fluent and disfluent trials compared to chance for monolinguals and bilinguals, in Blocks 1 and 2

	Monolinguals					Bilinguals				
	M	*SD*	*t*	*p*	*d*	*M*	*SD*	*t*	*p*	*d*
Raw Values										
Block 1										
Fluent	.48	.29	0.29	.78	0.07	.45	.25	– 0.74	.47	–0.19
Disfluent	.59	.22	1.65	.12	0.41	.60	.33	1.17	.26	0.29
Block 2										
Fluent	.44	.32	– 0.72	.48	–0.19	.55	.39	0.42	.69	0.13
Disfluent	.62	.27	1.74	.10	0.45	.44	.26	– 0.73	.48	–0.22
Baseline-Corrected Values										
Block 1										
Fluent	.03	.32	0.39	.70	0.10	.14	.30	1.86	.08†	0.47
Disfluent	.00	.32	0.05	.96	0.01	.08	.27	1.17	.26	0.29
Block 2										
Fluent	.11	.28	1.51	.15	0.38	-.13	.23	-2.07	.06†	-0.57
Disfluent	-.01	.42	-0.05	.96	-0.01	-.04	.28	-0.52	.62	-0.16

Note. *M* = Mean. *SD* = Standard deviation. †*p* < .10; **p* < .05; ***p* < .01. *t* values are comparisons to chance looking (chance = 0.5 for raw values before baseline looking corrections; chance = 0 for baseline-corrected values after baseline looking corrections).

Table 5

Learning phase analyses in the Disfluency window of looking to the novel object on fluent and disfluent trials compared to chance (0.5) for monolinguals and bilinguals, in Blocks 1 and 2

	Monolinguals					Bilinguals				
	M	*SD*	*t*	*p*	*d*	*M*	*SD*	*t*	*p*	*d*
Block 1										
Fluent	.51	.14	0.255	.80	0.06	.52	.19	0.509	.62	0.13
Disfluent	.49	.12	– 0.35	.73	–0.08	.41	.14	– 2.579	.02*	–0.64
Block 2										
Fluent	.37	.13	– 4.05	.001**	–1.01	.26	.20	– 4.71	< .001**	–1.18
Disfluent	.37	.25	– 2.12	.05*	–0.53	.32	.17	– 4.20	< .001**	–1.05

Note. *M* = Mean. *SD* = Standard deviation. †$p < .10$; *$p < .05$; **$p < .01$ *t* values are comparisons to chance looking (0.5).

Table 6

Learning phase analyses in the Label window of looking to the target object on fluent and disfluent trials compared to chance (0.5) for monolinguals and bilinguals, in Blocks 1 and 2

	Monolinguals					Bilinguals				
	M	*SD*	*t*	*p*	*d*	*M*	*SD*	*t*	*p*	*d*
Block 1										
Fluent	.70	.17	4.735	< .001**	1.15	.58	.21	1.621	.13	0.41
Disfluent	.65	.20	3.00	.008**	0.73	.67	.19	3.562	.002**	0.89
Block 2										
Fluent	.65	.15	4.190	< .001**	1.05	.59	.17	2.168	.05*	0.54
Disfluent	.68	.18	3.88	.001**	0.97	.60	.21	1.870	.08†	0.47

Note. *M* = Mean. *SD* = Standard deviation. †$p < .10$; *$p < .05$; **$p < .01$ *t* values are comparisons to chance looking (0.5).

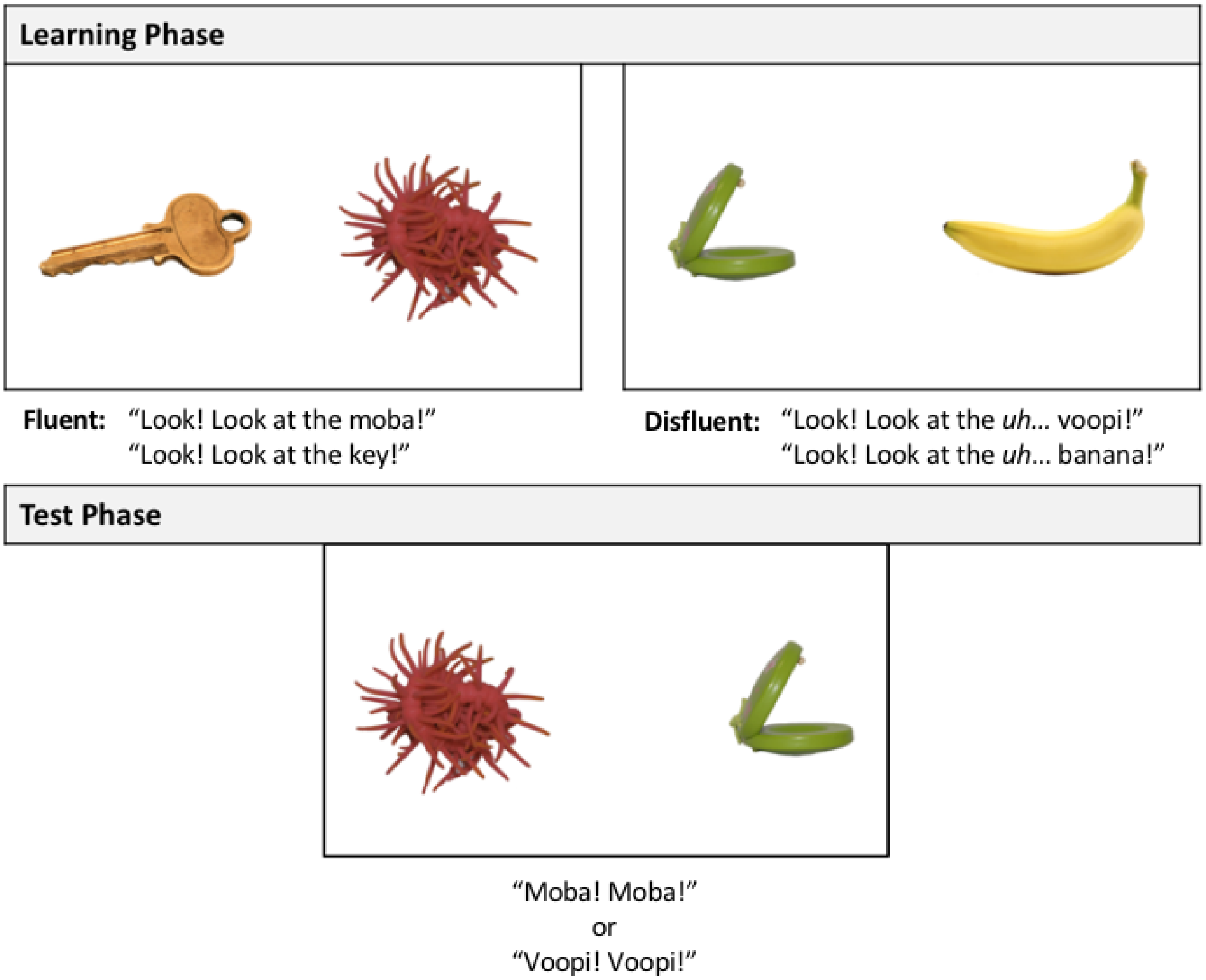


Figure 4. Manuscript 2: Visual of sample trials in the Learning phase and in the Test phase in English. Top: Sample Learning phase novel-familiar object pairings (Left: key-moba; Right: voopi-banana). Bottom: Sample Test phase novel-novel object pairing (moba-voopi).

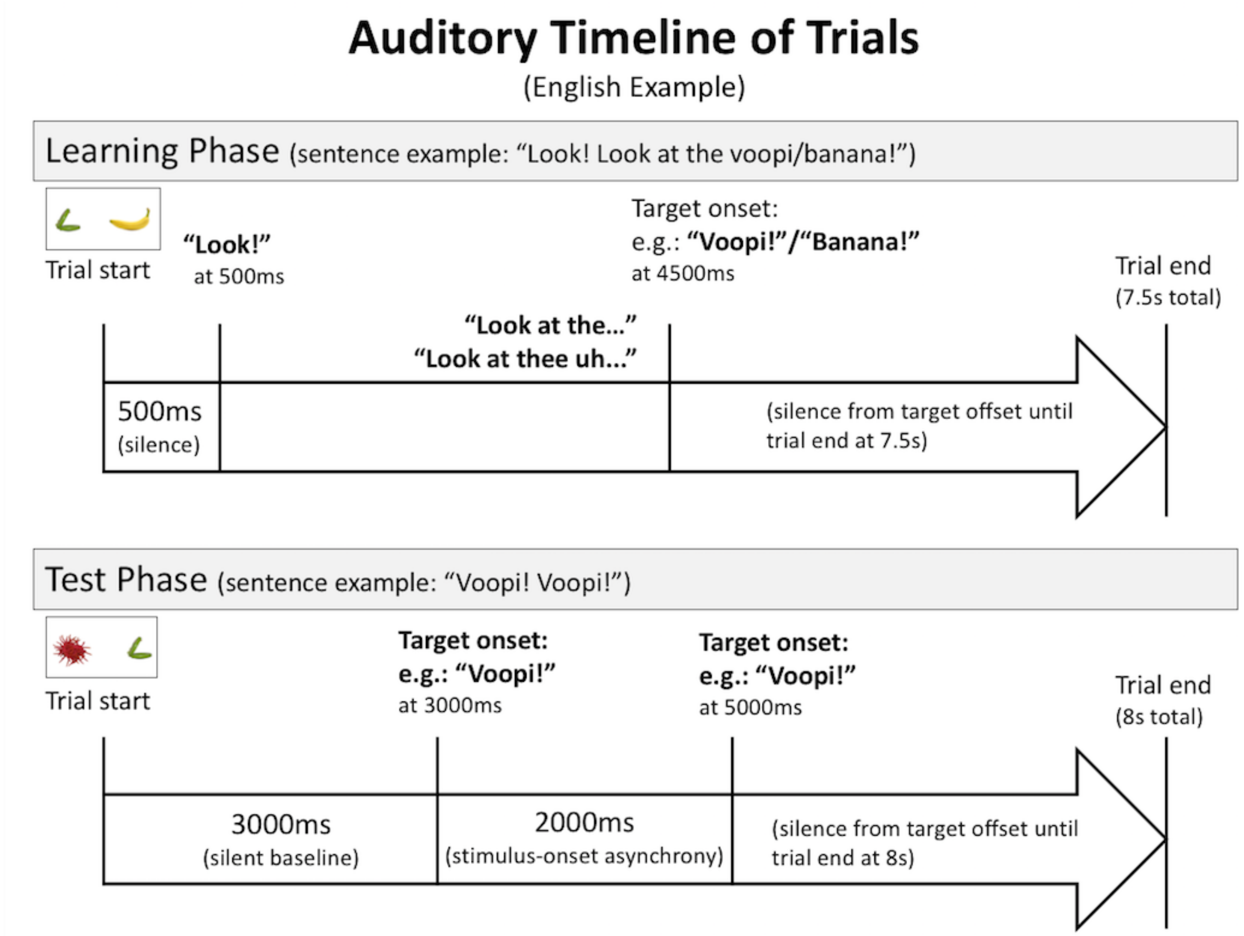


Figure 5. Manuscript 2: The auditory timeline of trials (English example) in the Learning Phase and in the Test Phase. Visual objects appeared on the screen at the start of each trial, and remained on the screen until the trial end. Stimulus-onset asynchrony refers to the constant time between the target offset and the next target onset (2000ms).

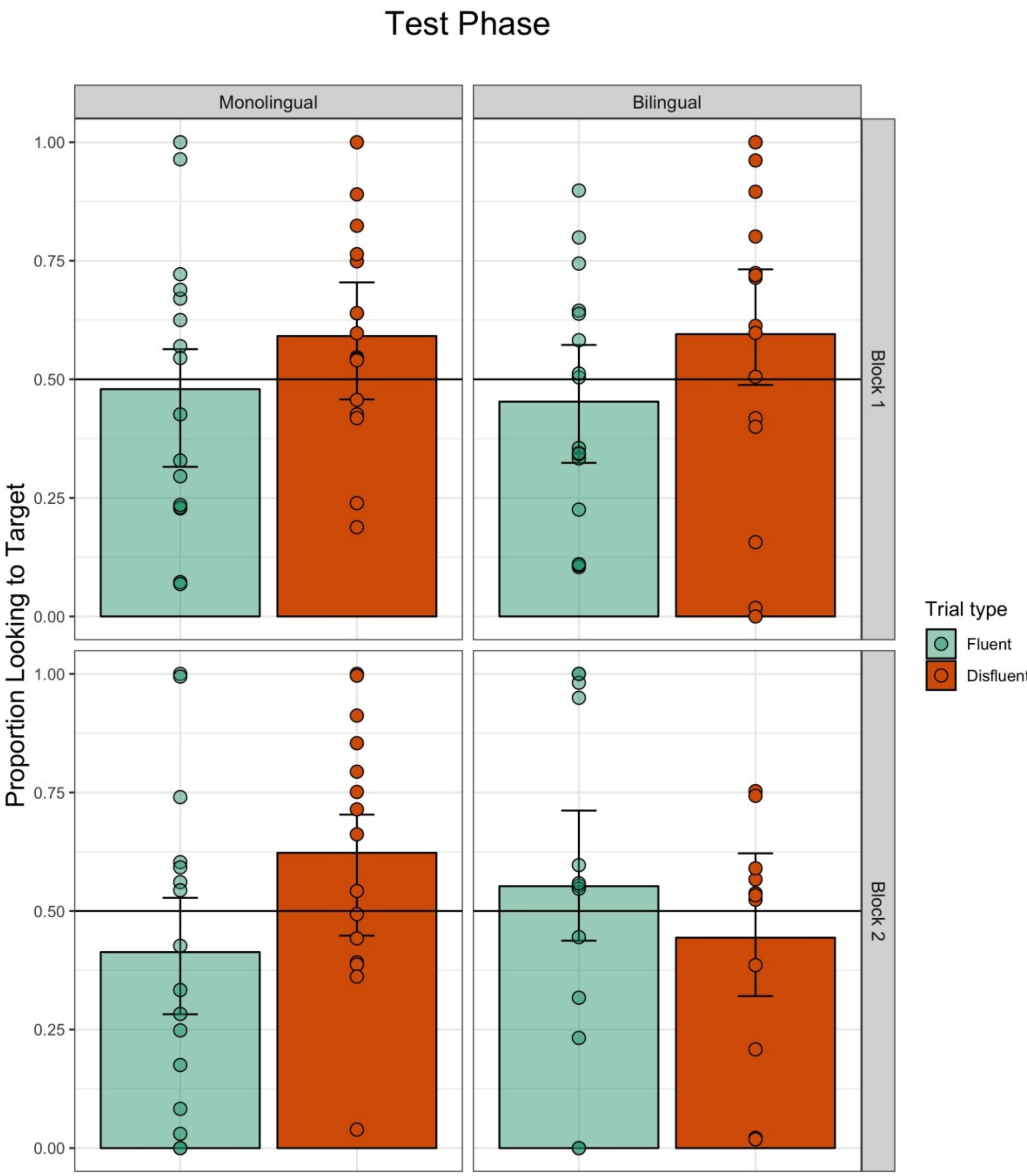


Figure 6. Manuscript 2: Looking to the target object on Fluent vs. Disfluent Trials in the Test phase, in Block 1 and Block 2. Chance looking (0.5) is represented by the horizontal line. Points show individual means. Error bars show the standard error.

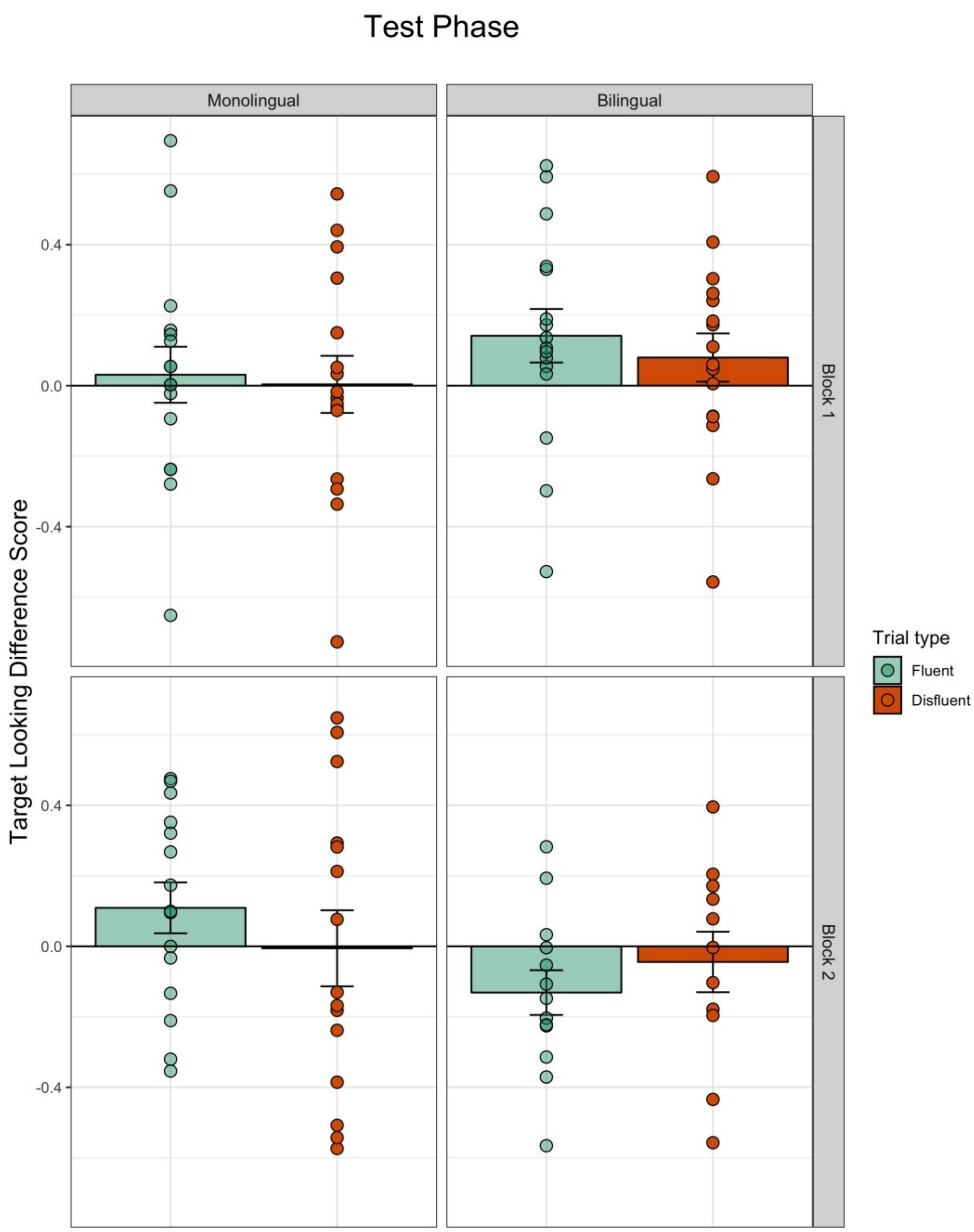


Figure 7. Manuscript 2: Baseline-corrected target looking in the Test phase, in Block 1 and Block 2. Chance looking (0) is represented by the horizontal line. Points show individual means. Error bars show the standard error of the mean. Positive difference scores correspond to an increase in looking to the target following the target label compared to the silent baseline, and negative difference scores correspond to a decrease in looking to the target following the target label compared to the silent baseline.

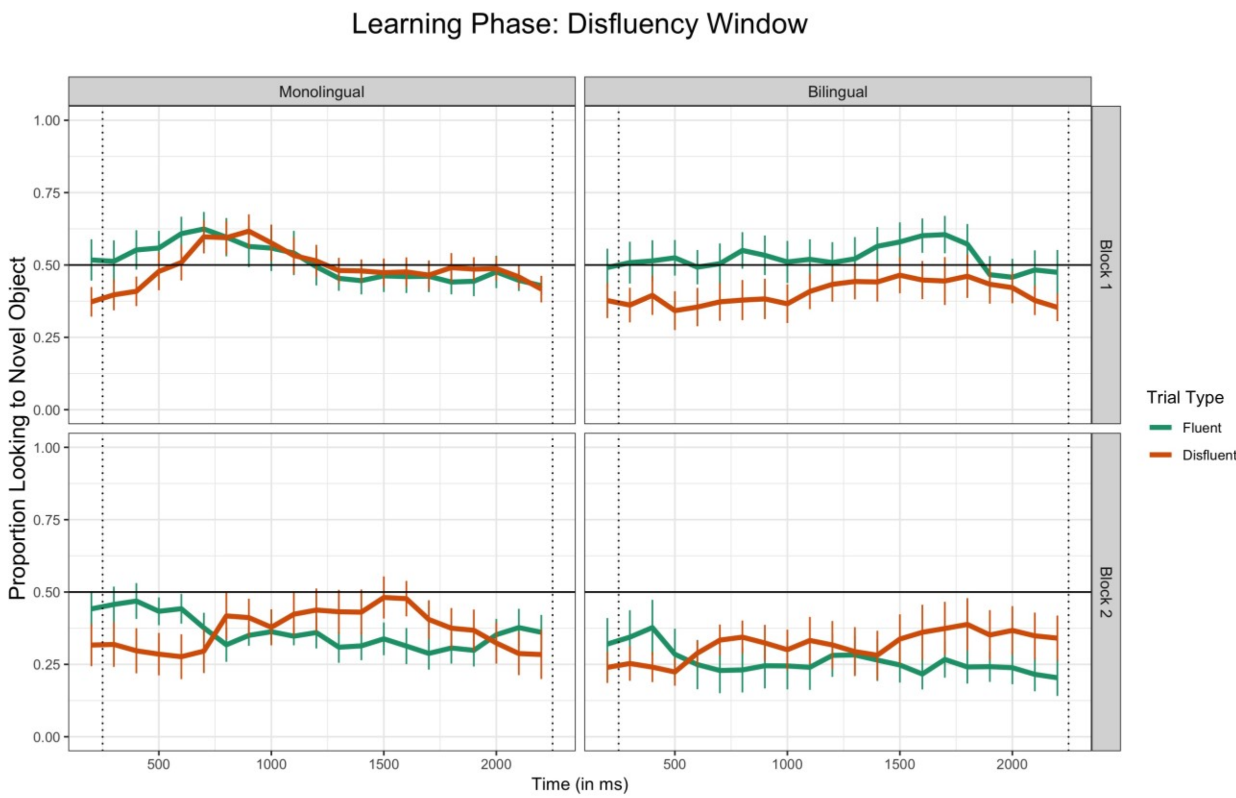


Figure 8. Manuscript 2: Timecourse of looking to the novel object in the Disfluency window in the Learning phase, in Block 1 and Block 2. Chance looking (0.5) is represented by the horizontal line. Vertical lines over timecourse data show the standard error of the mean.

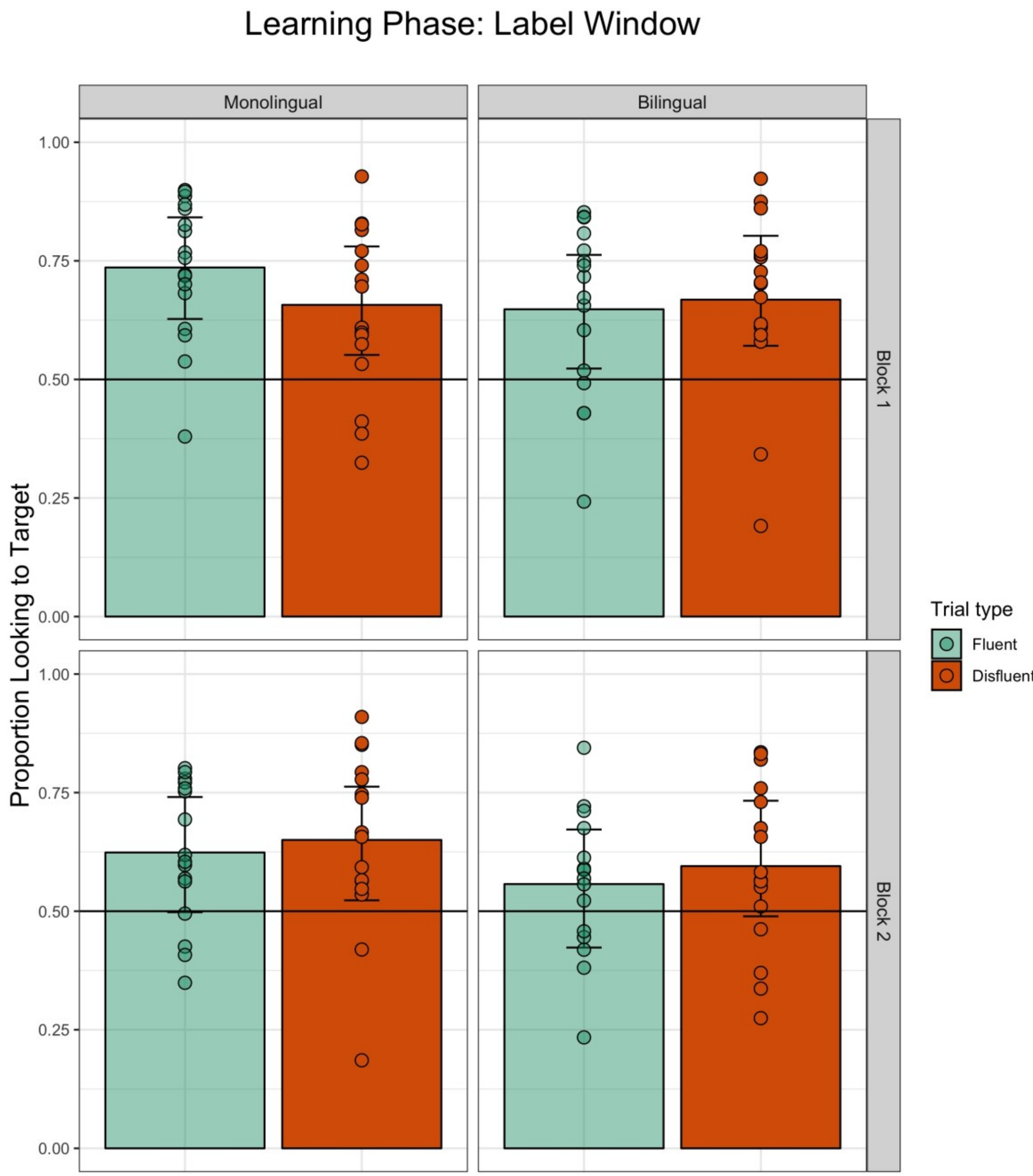


Figure 9. Manuscript 2: Looking to the target object in the Label window in the Learning phase, in Block 1 and Block 2. Chance looking (0.5) is represented by the horizontal line. Points show individual means. Error bars show the standard error.

4 CHAPTER 4: General Discussion

The overarching goal of the three studies reported in this dissertation was to understand the role of speech disfluencies in monolingual and bilingual children's language acquisition. Manuscript 1 (Chapter 2) investigated the predictive value of disfluencies in real-time speech comprehension. Can listeners from different language backgrounds use speech disfluencies that are non-native equally well as those that are native? We replicated and extended previous results, finding that both monolingual and bilingual toddlers, as well as bilingual adults, use speech disfluencies to predict that a speaker will label a novel rather than a familiar referent. Contrary to predictions, this pattern of results was unchanged whether toddlers were monolingual or bilingual, whether stimuli were in the dominant or non-dominant language, or whether speech disfluencies were native-sounding or non-native sounding to listeners.

Given the robustness of speech disfluencies for orienting children to a novel referent in Manuscript 1, Manuscript 2 (Chapter 3) investigated how this would affect word learning. Specifically, we asked whether disfluencies produced just before the labelling of novel words would affect word learning. We put forward two competing predictions: Either (1) the learning of the novel word would be enhanced if disfluencies direct children's attention to a novel object, or (2) the learning of the novel word would be hindered if children avoid learning from a disfluent, potentially hesitant speaker. Contrary to either prediction, we did not find robust differences between word learning in disfluent and fluent conditions in the test phase, even when considering the differing salience of novel objects. Surprisingly, during the learning phase, we did not replicate the pattern reported in Manuscript 1, as well as in the literature (Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016), that children increased their attention to a novel and/or discourse-new object upon hearing the disfluency. Although children did not appear to use the disfluency to direct their attention to the novel object, they seemed to look more to the familiar object regardless of whether a disfluency was heard. Importantly, children did look at the novel object upon hearing the novel word itself. Yet, despite identifying the intended referent, we found little evidence that children successfully made a word-object mapping. Thus, Manuscript 2 failed to find evidence of word learning in a context where a speaker was sometimes disfluent.

All together, these studies yielded an unpredicted pattern of results, suggesting that processing of speech disfluencies is more complex than anticipated. In the next few sections, I will discuss the main contributions, findings, and open questions raised by this research.

4.1 Main Contributions

As I have highlighted in the general introduction of this dissertation (Chapter 1), speech is full of disfluencies like *uh*, *euh*, and many others. Yet, very young children can navigate the intricacies of natural speech and accomplish the amazing feat of acquiring one, two, or more languages. What does hearing disfluencies in the speech of their caregivers mean for children acquiring their language(s)? This question sparked my interest in learning more about the role of disfluent speech in language acquisition.

The findings from the studies presented in this dissertation have addressed several important questions regarding the role of speech disfluencies in comprehension and word learning. The results clearly indicate that the language of the filled pause disfluency itself is not central to its predictive use. Manuscript 1 showed that children from monolingual and bilingual profiles, and bilingual adults, can use both native, non-native, and language-inconsistent disfluencies alike to predict that a speaker will label a novel object over a familiar one. This was a surprising finding, since there are several differences in monolingual and bilingual children's experience with language, and differences between the two groups on several dimensions have been found in previous studies, from mutual exclusivity in word comprehension (Halberda, 2003; Byers-Heinlein & Werker, 2009; 2013; Davidson, Jergovic, Imami, & Theodos, 1997; Davidson & Tell, 2005; Houston-Price, Caloghiris, & Raviglione, 2010) to cognitive tasks such as executive functioning (see review in Bialystok, 2001; Bialystok, Martin, & Viswanathan, 2005, though see Paap, Johnson, & Sawi, 2015). The fact that disfluencies from two different languages were used in the same way highlights the possibility that disfluencies may not be attributed to a specific language and may not be lexicalized in the same way that words are. Moreover, this finding reinforces the idea that both children of monolingual and of bilingual backgrounds can learn their languages successfully (Byers-Heinlein & Lew-Williams, 2013), and can use speech disfluencies from two languages during language comprehension, whether they are from their native language(s) or not.

While these studies have contributed significantly to our understanding of how children use disfluent speech during word recognition and word learning, it should be noted that the current findings are largely inconsistent with the original hypotheses. It was predicted that exposure to speech disfluencies in a specific language was important for listeners of different language backgrounds to use them predictively during real-time comprehension, and that listeners from different language backgrounds might learn novel words following speech disfluencies differently. Thus, much like how the experience with specific native-language sounds and words shape language processing (Kuhl et al., 2006; Werker & Tees, 1984; Werker, Yeung, & Yoshida, 2012), it was predicted that exposure to specific disfluencies in one's language(s) would influence how they were used in comprehension and for word learning. However, the data from both manuscripts was inconsistent with this original hypothesis, instead, favouring a larger role for social factors or characteristics of disfluencies common across languages.

The results in this dissertation suggest that exposure to a range of speakers who are disfluent may be more important in shaping the disfluency effect than language-specific exposure to speech disfluencies. This finding is particularly relevant for children who may be exposed to non-native speakers who produce disfluencies that are non-native to children, as children may still be able to predict the contents of that speaker's discourse using their disfluencies. Thus, the present research does not support the position that the language or specific phonetic realization of the filled pause disfluency (e.g., the English *uh* vs. the French *euh*) affects comprehension or the predictions that children make about upcoming speech.

It is important to note that children's speech processing is affected by subtle phonetic changes and language switches, even though it was not in the case of disfluencies. For example, previous research shows that vowel mispronunciations (e.g., pez (fish) vs. pɛz (fish mispronounced)) can impair word recognition in Catalan-Spanish bilingual infants aged about 18 to 26 months, though only when the vowels substituted are contrastive in the child's language (i.e., a change in a vowel could change the meaning of a word; Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009). Similarly, a switch in languages at the word level (e.g., "Find the chien!" vs. "Find the dog!"), has also been found to challenge processing during comprehension in English-French and English-Spanish bilingual children aged 20 months and 18-30 months, respectively (Byers-Heinlein, Morin-Lessard, & Lew-Williams, 2017; Potter, Fourakis, Morin-Lessard, Byers-Heinlein, & Lew-Williams, *under review*; although see Jardak, Lew-Williams, & Byers-Heinlein, *in prep*., for a case when comprehension as unaffected by language switching). It will be important for future research to understand whether the processing of switches in non-lexical parts of speech, such as speech disfluencies, are processed differently from switches that occur in the context of words.

Speech disfluencies, whether with a switch in phonetics or not, can be used by listeners to predict upcoming information during language processing. Yet, even when listeners do make use of speech disfluencies during language processing, this information does not override lexical information that follows it. That is, when hearing "Look at the *uh*… dog!", listeners might initially look at a novel object when hearing *uh* but then quickly shift their gaze to the target object dog when they hear the noun (e.g., as in Manuscript 1). When listeners do not use speech disfluencies to predict a novel referent during processing, they can ultimately look at the correct target (e.g., as in Manuscript 2). This finding of correct looking to the target was obtained in both Manuscript 1 and Manuscript 2. In Manuscript 1, which focused on the comprehension of disfluencies, we found that while bilingual children and adults were more accurate at identifying familiar objects compared to novel objects, looking was nonetheless above chance for all children, and for both novel and familiar objects. In Manuscript 2, which focused on the learning of words following disfluencies, children did not show the disfluency effect and failed to show strong evidence of novel word learning. However, children were nonetheless able to identify the target object once it was labeled in the learning phase, suggesting that disfluencies do not disrupt listeners' ability to ultimately identify a word labeled by a speaker. For example, upon hearing "Look at the uh… X!", listeners ultimately looked at the appropriate labeled object, whether this object is familiar or novel. Overall, we found that disfluencies do not disrupt comprehension of familiar words that follow them.

Another important discovery is that there may be a dissociation between children's use of disfluencies for identifying a speaker's intended referent (as found in Manuscript 1), and using this information to learn a novel word (as seen in Manuscript 2). Indeed, while the results of Manuscript 1 showed evidence of predictive looking to the novel object upon hearing a disfluency, the results from Manuscript 2 indicated that disfluencies do not necessarily promote word learning. Indeed, disfluencies may not be advantageous for – and potentially even hinder – word learning, particularly if children see them as a marker of a speaker's lack of competence. Thus, the flipside of disfluencies may be that for word learning, they are not particularly useful. Instead of facilitating word learning, disfluencies may affect children's willingness to learn novel words, especially if this willingness was dependent on the belief that they can rely on a speaker. Since novel word learning in fluent conditions may already be challenging for young learners, having to consider a speaker's state and knowledge may prove too cognitively demanding for them to succeed on a word learning task. Although more research is needed before coming to stronger conclusions on the effect of disfluencies on word learning, the results from Manuscript 2 are consistent with the possibility that disfluencies may prevent children from learning novel words.

Overall, the discoveries made in this dissertation seem to converge in an important way. A speaker's knowledge, and listeners' perception of a speaker's intent might be important for how disfluencies are viewed and used by listeners. Rather than being anchored in factors specific to a listeners' native language(s), disfluencies may be rooted in characteristics that are shared across different languages (e.g., lengthening, suggesting uncertainty from a speaker), and are therefore useful for listeners of different language backgrounds. Moreover, the presence of disfluent speech in utterances does not seem to affect listener's comprehension of a message. However, disfluencies may prevent learning if a speaker is perceived to be uncertain about a novel object's label in the context of word learning. This leaves open the possibility that experience hearing disfluent speakers naming novel or discourse-new referents over time signals a speaker's uncertainty for listeners. The next section will provide a more in-depth discussion of the social nature of speech disfluencies, based on the findings in this dissertation and previous research.

4.2 Speech Disfluencies as a Social Cue

Overwhelmingly, this dissertation, as well as studies that have been published since this research was originally conceptualized (Orena & White, 2015; Owens & Graham, 2016) point to certain general aspects of speakers or disfluencies that may be useful for children to form predictions about upcoming speech. As discussed in the general introduction of this dissertation, speech disfluencies appear in predictable locations in the speech stream, often prior to a novel or discourse-new word. However, this systematicity is also related to a deeper reason connected to speakers themselves. That is, speakers tend to be less certain, or to experience retrieval difficulty before producing novel or discourse-new words.

Our findings are consistent with the position that children are sensitive to deeper, speaker-related causes of disfluencies, and are not simply tracking co-occurrences between disfluencies and novelty. In the comprehension study in Manuscript 1, this was evidenced by the fact that monolinguals and bilinguals, who have different specific experiences hearing specific realizations of disfluencies followed by novel or discourse-new words, can use atypical non-native (non-dominant in the case of bilinguals) disfluencies as well as native disfluencies one to predict a speaker's intention to label a novel object. In the word learning study in Manuscript 2, we found little evidence that children learned any words, either presented after speech disfluencies or presented in fluent speech. This could indicate that disfluencies hinder word learning, and that children understand that a speaker who is uncertain may not be a good source of information for word learning. For example, when choosing whether to learn language from speakers, children might rely on whether they believe that someone is confident or not. Lack of confidence may hint to a speaker's potential inaccuracy in their labelling of an object, and research has shown that children are reluctant to imitate unconfident speakers (Brosseau-Liard & Poulin-Dubois, 2014; Brosseau-Liard & Poulin-Dubois, 2016) and learn from previously inaccurate informants (Brooker & Poulin-Dubois, 2013; Brosseau-Liard, Penney, & Poulin-Dubois, 2015). Whether the speaker is actually accurate in their labelling of an object may not matter, as this speaker's hesitations may be sufficient for children to experience the uneasy feeling that the speaker is not *that* knowledgeable. Prior research on the predictive use of disfluencies points to children considering a speaker's knowledge state, specifically whether a speaker is forgetful or not, when using disfluencies as a cue to a speaker's referential intent (Orena & White, 2015). Although children may use disfluencies predictively when a speaker is explicitly knowledgeable during comprehension, they may require more confidence from a speaker before committing to learning.

Children's perceived uncertainty from a speaker may be exacerbated when cues about this speaker are reduced. In eyetracking studies like the ones presented in this dissertation, there is typically no visible speaker present to accompany the speech that children hear. This means that cues about a speaker's confidence are diminished and are limited to the speaker's utterances, and that children's judgements of a speaker will be based on these cues only. This single source of information on a speaker may be less forgiving than studies with a live speaker, as some ostensive cues provided by speakers that are present in live interactions (e.g., eye gaze, contingent actions, posture) may reveal more confidence, and in turn be more conducive to the learning of words. For example, the fact that the speaker is a trained research assistant and is in a laboratory setting may be sufficient for children to attribute more confidence to that speaker despite this speaker's disfluencies, compared to if children only heard this speaker. Below, I discuss how modality and other factors may play a role in the disfluency effect and children's confidence and learning of words.

4.2.1 Factors influencing the disfluency effect

One puzzling finding from the research findings reported in this dissertation was the failure to replicate the disfluency effect (i.e., greater looking at a novel object upon hearing a novel word) in Manuscript 2, which tested word learning. This was surprising since we observed a disfluency effect across-the-board in Manuscript 1. Specifically, in Manuscript 2, children did not look more to the novel object than the familiar object upon hearing the disfluency in the Learning phase, instead looking more to the familiar object regardless of whether they heard a disfluency or not. This was a surprising finding, given that the disfluency effect was observed quite robustly across children and adults, and across listeners of different language backgrounds in Manuscript 1. Moreover, the disfluency effect has also been observed in studies with children of the same age and older than children in the current study (Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016).

There were, however, notable differences between the study design of Manuscript 2 and the design of the studies in Manuscript 1. Compared to Manuscript 2, Manuscript 1 was more similar in experimental design to previous studies with children, in terms of the number, type, structure, and variety of trials (Kidd et al., 2012; Owens & Graham, 2016; Orena & White, 2015). The first important difference between the design of Manuscript 1 and Manuscript 2 and other studies is that, in Manuscript 2, the status of the novel word as new in the discourse had been removed. For example, previous studies always introduced the familiar object twice before labelling the novel object, either via labelling of the familiar object (e.g., "I see the ball!", "Ooh, what a nice ball!"; Kidd et al., 2012; Owens & Graham, 2016) or via looking to the novel object (e.g., a puppet always looked at the cookie as opposed to the blimmick, a novel object; Orena & White, 2015). Indeed, Manuscript 1 and previous studies with children (Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016) have found a disfluency effect when a familiar object had been previously identified in the discourse, making the object introduced in the third presentation stand out as new to the discourse. In fact, a recent study found no disfluency effect in children when prior discourse information was removed, though it was found in adults who based their predictions on object novelty alone (Owens, Thacker, & Graham, 2017). Other research has also indicated that prior discourse is often informative for word learning (Fernald, Frank, Goodman, & Tenenbaum, 2009). In Manuscript 2 of this dissertation, presenting the novel object in a single sentence, without establishing discourse context, was a deliberate choice made to simplify the design of the word learning study, and to reduce the overall length of the study so that children would remain attentive. This design decision may have inadvertently deprived children of some important discourse information that they use when using disfluencies predictively, and the brief exposure to the familiar object may have resulted in children looking more to the familiar object during the disfluency in some cases (Hunter, Ross, & Ames, 1982). Although other factors may have contributed to the failure of finding a disfluency effect in Manuscript 2, this change in our design provides additional evidence that prior discourse establishing a novel object as new to the discourse is a driver of the disfluency effect in children (Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016; Owens, Thacker, & Graham, 2017).

Another aspect of the design in Manuscript 2 that might have affected children's performance is the repeated presentation of the same two novel words in the learning phase. Indeed, instead of presenting children with a different novel on each trial as was done in Manuscript 1 and in previous studies (Kidd, White, & Aslin, 2011; Orena & White, 2015), the same two novel words were presented on multiple occasions in Manuscript 2. Given that the focus of Manuscript 2 was novel word learning, this repeated presentation was essential to give children enough opportunities to learn the novel words. However, these multiple exposures to the novel words may have affected the novelty of the words over the course of the study (Samuelson & Smith, 1998), which could have disrupted the disfluency effect. For example, children might have recalled that the novel word 'moba' and its referent object was introduced in earlier trials, making it more familiar as they continued to see it presented over time.

Finally, the unequal salience of the two novel objects in Manuscript 2 may also have affected children's performance at test. Because of other manipulations required for the word learning study, although side of presentation for each object was counterbalanced, each novel label was always paired with the same novel object. While the two novel objects had been judged by adults to be of similar salience, children's looking behavior in a silent baseline showed that they found the "moba" more visually interesting than the "voopi". Because the two objects were always presented side-by-side during the test phase, this difference in salience might have masked word learning. Our findings suggest that object salience is not only a relevant consideration when assessing word learning in infants (Hollich et al., 2000), but it remains so even in toddler-aged children who are older than ages at which children have been found to ignore salience cues for other cues (e.g., social cues; Pruden et al., 2006). This salience likely influenced children's looks, making it difficult for us to draw stronger conclusions of word learning in the context of speech disfluencies.

Together, these design differences highlight that the disfluency effect can be affected by the removal of the prior discourse establishing the novel object as new to the discourse, by the repetitive nature of word learning studies, the salience of objects, or a combination of some or all of the three. Therefore, the two manuscripts from this study – and design comparisons with previous research – suggest that for disfluencies to potentially be beneficial during comprehension, some characteristics of the discourse (e.g., discourse novelty) and of the novel object need to be present.

4.2.2 The impact of disfluencies on word learning

Previous work on children's processing of disfluencies has focused on their use in language comprehension, with little empirical attention paid to word learning. The present work from Manuscript 2 on the role of disfluencies in word learning is important because based on previous comprehension studies that found a disfluency effect in children (Manuscript 1 of this dissertation, and Kidd, White, & Aslin, 2011; Orena & White, 2015; Owens & Graham, 2016), it would seem that the increased attention to the novel object might support word learning. However, no previous research had actually tested children's word learning in the context of disfluencies. In the word learning study in Manuscript 2, we found that children did not learn any of the novel words, even the novel words presented in fluent trials. This was surprising, given that children at this age and much younger (by age 14 months) can successfully learn words in similar looking paradigms with a similar number of learning trials (Byers-Heinlein, Fennell, & Werker, 2013; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Potentially, this happened due to the same factors that may have disrupted the disfluency effect, namely the removal of the discourse-novelty of the novel object from Manuscript 1 to Manuscript 2, the repeated presentations of the novel word over the course of the study, or a combination the two.

Another possibility is that the disfluency itself made the learning of the novel word following it more difficult, and this difficulty generalized to the fluent condition. If this is the case, this would support the hypothesis that speech disfluencies hinder learning of novel words that follow them. However, it is difficult to draw a strong conclusion since we did not observe any word learning in either the fluent or the disfluent condition. We did note a pattern of the means being larger on the disfluent trials than on the fluent trials in three instances out of four (for monolinguals in blocks 1 and 2, and for bilinguals in block 1). Nonetheless, this pattern did not reach statistical significance, and must be interpreted with caution given that a smaller number of participants contributed data over the course of the test phase, and given the lack of the disfluency effect found in the learning phase. The follow-up condition suggested in Manuscript 2 – with fluent-only trials – may help illuminate whether or not disfluencies in our study made it too challenging for children to learn words. Another way to determine how much disfluencies in our study impacted word learning would be to investigate whether silences of the same length of the disfluencies used in our study affect word learning compared to linguistic disfluencies (i.e., *uh* and *euh*). Silences are attributed by adult listeners as signaling a greater production difficulty compared to filled pauses such as *um* (Brennan & Williams, 1995), since the presence of filled pauses indicates that speakers can anticipate a certain delay compared to when they pause silently (Fox Tree, 2001). Finding that these non-linguistic speech disfluencies do not affect word learning – whereas linguistic disfluencies do – would provide further evidence that listeners are interpreting filled pauses as representing a speaker's uncertainty, either on their own or via characteristics of the filled pauses themselves (e.g., lengthening of a vowel sound). Further, it would highlight that compared to other breaks in speech, speech disfluencies such as filled pauses carry a distinctive social quality that listeners take into consideration when learning words.

However, more recent research conducted since the conceptualization of this dissertation has indicated that children sometimes do learn new words in the context of disfluencies. One study recently investigated word learning following speech disfluencies with 3- and 4-year old children (White, Nilsen, & Riemersma, *under review*). Instead of teaching children new labels using recorded stimuli, this study investigated word learning in a live interaction study with a speaker. The speaker produced a disfluency prior to naming some objects, and was fluent prior to naming some others. Following this live interaction, children's word learning was then assessed using preferential looking, with two objects on a screen being either both familiar (to maintain children's attention) or both novel (to assess word learning). Unlike our study in Manuscript 2 in which children were taught only two novel words, four novel words were taught in the live interaction study. Since two words of each type were taught (2 fluent, 2 disfluent), this allowed for the presentation of pairs of novel objects of the same type at test. The results contrasted with our word learning study results in Manuscript 2: Children aged 3- and 4-years both successfully learned the novel words, and learned them equally well in the fluent and disfluent conditions.

There are several differences between this emerging research and the study reported in Manuscript 2. The first obvious one is the modality. As discussed above, eyetracking studies like our design, in which there is no speaker present, reduce cues present in live interactions that may facilitate word learning. If social cues are reduced, listeners' judgements of a speaker are based on these cues only. This may result in listeners making more negative judgements about the speakers' own personal characteristics (in this case lack of knowledge) given that they are deprived of other potentially informative cues (e.g., ostensive cues) that could explain why a speaker might be disfluent (Brooks & Meltzoff, 2002; Colonnesi, Stams, Koster, & Noom, 2010; Emery, 2000; Vignovic & Thompson, 2010; Yow & Markman, 2016). A second difference worth noting is the difference in ages between children in the current study and in the live interaction study. Indeed, children in our study were aged 32 months (2 years 8 months), a few months younger than the 3-year-olds tested in the live interaction study (White, Nilsen, & Riemersma, *under review*). It is possible that the presence of disfluencies poses a greater problem for word learning in children prior to age 3, and that word learning in the context of disfluencies is an ability that appears later in development. Finally, a third difference that may affect children's learning in the context of disfluent speech is its perceived naturalness, which may partly be established by the different lengths of the disfluency. Indeed, the disfluencies produced by the speaker in the live study were shorter (810 ms on average) than we used in the studies in this dissertation, specifically in Manuscript 2. In English, the filled pause lasted 916 ms on average (*uh*), and extended to 2195 ms when measured with the lengthened determiner (*thee uh*). In French, the filled pause lasted 953 ms on average (*euh*) and extended to 2334 ms on average with the lengthened determiner (*la euh*). The length of disfluencies produced in the live study was also shorter than that used in previous disfluencies studies with children (e.g., 1130 ms (*uh*) and 1600 ms (*thee uh*) in Orena & White, 2015). While longer disfluencies may be helpful for children during comprehension, such longer disfluencies may be more problematic for word learning compared to shorter ones, as they may indicate a greater uncertainty on the part of speakers. This possibility would need to be investigated in future research. Together, findings from the word learning study in Manuscript 2 and from the live interaction study (White, Nilsen, & Riemersma, *under review*) seem to indicate that disfluencies themselves may not completely prevent children from learning new words. However, one or several components of our design in Manuscript 2 made it particularly challenging for children to learn the novel words. The particular demands of the task, children's age, and characteristics of the disfluencies themselves are all features that will be important to consider and further investigate in future research.

Ultimately, irrespective of speech disfluencies, we did not find evidence of retention in the word learning study. This retention aspect of word learning is generally something that can be challenging for toddlers (Bion, Borovsky, & Fernald, 2013; Horst & Samuelson, 2008). For example, in our task, for successful word learning, children first needed to identify the novel object on the visual display using their knowledge of the familiar object, and infer that the unknown object was the most likely referent. Then, children would have had to access the mental representation of the novel object and the novel word and link the two together, and do this for each of the two novel words. To show word learning at test, children would have had to access their freshly-established knowledge of each novel word in order to infer which of the two novel objects was the correct referent, and overcome differences in objects' visual salience. Adding to this uncertainty from the speaker may have added an extra layer of difficulty to an already difficult cognitive task for children. The fluent-only condition follow-up study will be important in telling us which factors may contribute or not to the lack of word learning observed in Manuscript 2 of this dissertation.

4.3 Speech Disfluencies: Do they support Language Acquisition in Young Learners?

A question of practical importance stemming from this dissertation is whether disfluencies are beneficial, neutral, or hindering language acquisition, beyond the question of word learning and comprehension. In this section, I will entertain each of these possibilities in turn.

Let us first discuss how speech disfluencies could be considered beneficial. Speech disfluencies are a normal feature of spontaneous conversation and of natural spoken language: They are produced in different languages (Duez, 1982; Fischer, 2000), often in predictable locations (Clark & Wasow, 1998; Shriberg, 1996), and generally denote a certain difficulty in language planning and production (Engelhardt, Nigg, & Ferreira, 2013; Fox Tree & Clark, 1997; Shriberg, 1996; Vitkovitch & Tyrrell, 1995). For speakers, disfluencies serve a range of communicative purposes, including to remain relatively fluent and to express a production difficulty to listeners (Clark & Wilkes-Gibbs, 1990), to name a few. For listeners, disfluencies can also be a useful tool. We saw in Manuscript 1 on real-time comprehension that they can be useful for children and adults to make predictions about an upcoming word. Though statistical results from Manuscript 2 would not warrant the making of a strong claim that disfluencies may facilitate word learning, results nonetheless suggest that looking to the target object was greater on disfluent trials compared to on fluent trials, indicating a potential benefit of disfluencies over fluent speech, which will need to be explored further. Moreover, we found in Manuscript 1 that the specific realization of speech disfluencies does not seem central to their use. This is an important finding, as children learn their languages listening to speech, and speech is filled with speech disfluencies. In addition, children – bilinguals in particular – may hear disfluencies from different speakers in different languages, though our research from Manuscript 1 indicates that this is not problematic for listeners. The vowel change between the English *uh* and the French *euh* in Manuscript 1 did not affect the disfluency effect, and did not prevent listeners from correctly identifying the target word. In this respect, the evidence on the comprehension of words following speech disfluencies points to them being quite a helpful tool, which is freely available in speech.

Disfluencies may also be neutral in some instances. While they can be beneficial during real-time comprehension (Manuscript 1), they may not be particularly helpful or hindering during word learning. For instance, they may not make the learning of words better, or worse than if children heard fluent speech. Since disfluencies are frequent in speech, young listeners may simply come to discard disfluencies in a word learning situation, when other cues from the environment (e.g., objects) and heuristics (e.g., mutual exclusivity) are available to facilitate word learning. Given the inconclusive findings from Manuscript 2, children's failure to learn words may not be related to speech disfluencies, but rather, to other aspects of the design discussed in section 4.2.1 above.

Now, what if disfluencies are in fact not beneficial, and possibly detrimental for language acquisition? While emerging research suggests that children can learn words following disfluencies in live interactions (White, Nilsen, & Riemersma, *under review*), we did not observe any strong evidence of word learning in Manuscript 2. As pointed out above, there are several reasons why this could be. One of them was that the lengthened determiner in the live study was shorter (White, Nilsen, & Riemersma, *under review*) compared to the one used in Manuscript 2 and in other previous studies (Kidd et al., 2012; Orena & White, 2015; Owens & Graham, 2016). The length of disfluencies tends to vary, and this variation could be due to individual variability between speakers or based on the production difficulty experienced. If some children are exposed to speakers and caregivers who produce particularly long disfluencies, would they have a greater difficulty learning words? More specifically, might children attribute a greater difficulty and uncertainty to speakers who produce longer disfluencies in the context of word learning, even if this was not the source of the disfluency? This is a possibility based on the findings in Manuscript 2 on word learning.

Another reason why disfluencies may not be a particularly reliable cue to use for language acquisition, is that their use tends to change based on usage and speakers. Indeed, research has documented a change over time in the prevalence of *um* vs. *uh*, indicating that *um* is being produced more and more frequently in spontaneous speech. Younger speakers and females (Acton, 2011; Wieling et al., 2016) tend to use more *ums* compared to older speakers and men (Le Grezause, 2017). This is consistent with the general finding that younger speakers and female speakers often drive linguistic change (Labov, 1990), which could mean that *ums* may eventually be produced more frequently than *uhs*. Although research suggests that *um* tends to be produced before a greater disruption of speech compared to *uh* (Clark & Fox Tree, 2002), *um* may not signal greater speech disruptions if it is produced increasingly frequently and in place of *uh*. As language continues to evolve, an important question is whether speech disfluencies such as *uh* and *euh* will remain as useful as to listeners compared to *um* and *eum* (in French). If speech disfluencies are mentally represented based on their acoustic properties such as lengthening, various kinds of disfluencies may remain helpful, at least for comprehension.

Since the rate of use of different speech disfluencies can be influenced by one's peers and trends in linguistic productions, this raises the possibility that particular forms of disfluencies within a language are represented separately. On one hand, the research in this dissertation suggests that disfluencies may not be stored in memory the same way that lexical items are. Indeed, we found no difference in the predictive use of disfluencies that were in English versus in French, specifically *uh* and *euh*. However, it is quite likely that the particular forms of disfluencies such as *uh* and *um* are stored and mentally represented separately to some degree. This possibility is quite plausible, particularly if young people and female speakers produce one form over the other in spontaneous speech (Acton, 2011; Wieling et al., 2016). Specifically, if the use of *um* is becoming more frequent than *uh* for a particular social group, listeners would have to adapt their mental representations of when the use of one or the other is preferred based on what they hear in their environment. More research following natural data closely is needed to determine whether speech disfluencies will remain as informative and beneficial as their use changes over time and between different social groups.

In sum, speech disfluencies may support language acquisition in young learners in some instances, and possibly not in others. Future studies will allow us to gain a better understanding of whether the beneficial aspects of disfluencies in real-time comprehension outweigh the potentially not so beneficial aspects during word learning and language acquisition more generally.

4.4 Limitations and Future Directions

The studies presented here were pioneering, in that they included the first studies to investigate disfluency use in bilingual toddlers, as well as one of the first studies to specifically test word learning in the context of speech disfluencies. However, they have several limitations. These limitations have generated new research avenues and highlighted important factors to consider for follow-up studies on the role of speech disfluencies in comprehension and word learning. I will discuss these factors and considerations following each limitation.

A first limitation, which is particularly important in the context of Manuscript 1, is that the disfluencies produced in natural speech have several, correlated parts (e.g., the lengthening of the determiner and of the filled pause disfluency itself), and the current study was not designed to disentangle their separate effects. Although in this dissertation I have generally been referring to the filled pauses *uh* and *euh* more generally as "speech disfluencies" or simply "disfluencies", these tokens often do not appear on their own in natural speech. In fact, filled pauses are often preceded by a lengthened determiner (e.g., theee in English, *leee/laaa* in French*),* possibly with a vowel change (e.g., in English, a lengthened the is pronounced *thee* as opposed to *thuh*; Fox Tree & Clark, 1997), and vary in how much they are lengthened. In both studies, we chose to always use a lengthened determiner prior to the disfluency, in order to maximize ecological validity (i.e., this is what speakers typically do), and to stick closely to the methods used in previous research with children (Kidd et al., 2011; Orena & White, 2015). Had we found an effect of language-consistent (e.g., "Look at thee *uh*…") but not language-inconsistent (e.g., "Look at thee *euh*…")

disfluencies, we could have concluded that the form of the disfluency itself affects their predictive use. However, given we did not find an effect of this manipulation, an alternate explanation is that the lengthened determiner (which was present for all disfluencies) overshadowed such an effect. Thus, because of all the characteristics packaged with and around filled pauses that contribute to making speech disfluent, it is not possible at this stage to know how much these factors played a role in the disfluency effect observed. It will be important for future research to further explore the role of lengthening, without filled pauses (as well as the inverse), in the disfluency effect.

A second limitation is linked to the main finding that speaker-based characteristics may be at play, alone or together with general characteristics such as the lengthened determiner. While we have ruled out that specific associations between a novel or discourse-new object and the filled-pauses in a particular language, such as *uh* and *euh*, matter for the predictive use of speech disfluencies, it is unclear what general factors are truly involved. We have identified several factors that could play a role, for example speaker uncertainty and determiner lengthening, but more research is needed to pinpoint which ones are the main drivers of the disfluency effect in young listeners. Some research has found that the discourse novelty of the novel object is one such driver of the disfluency effect, at least in children (Owens & Graham, 2016; Owens, Thacker, & Graham, 2017). The role of the lengthened determiner that has been present in most studies investigating the comprehension of disfluencies in children (Kidd et al., 2011; Orena & White, 2015; Owens & Graham, 2016; Owens, Thacker, & Graham, 2017; Thacker, Chambers, & Graham, 2018) and adults (Arnold et al., 2003; Arnold et al., 2007; Arnold, Tanenhaus, Altmann, & Fagnano, 2004; Thacker, Chambers, & Graham, 2018) has yet to further be investigated. Manipulating other cues of speaker certainty and speaker knowledge in the context of disfluencies will be an important avenue of future research.

A third important limitation concerns the inconclusive results from the word learning study from Manuscript 2. Our study used a within-subject manipulation to investigate word learning in fluent versus disfluent sentences. Unexpectedly, we found no convincing evidence of word learning in either condition. In section 4.2.2 above on the impact of disfluencies on word learning, we have highlighted several factors that might have hindered word learning in our study design. However, it is difficult to interpret this type of null result, and to make firm conclusions about which factors in this study most affected word learning in Manuscript 2.

Further research, including our own follow-up study with fluent-only sentences, will help us elucidate which factors can explain the lack of evidence of word learning in children. If children are able to learn words in the same experimental design with fluent-only speech, it would mean that speech disfluencies can indeed hinder word learning, and that characteristics of the design (e.g., repeated presentations of the novel word, object salience) were not behind children's failure to learn any novel words. Nonetheless, our analysis of the learning phase rules out a problem in identifying the target referent.

In sum, future work should aim to disentangle which particular aspects of disfluencies drive the disfluency effect, and whether this effect is based on characteristics of the speaker or more broadly to exposure to speech disfluencies followed by a novel or discourse-new word.

4.5 Conclusions

This dissertation began with the idea that speech recognition devices use pre-established formulas to solve the problem of being presented with disfluent natural speech. Unlike these devices which are programmed with algorithms to handle imperfect speech, children have to solve the puzzle of language – which includes disfluent speech – as they go, and flexibly adapt their predictions based on cues from the visual environment, cues in the speech stream, cues from speakers, and their own pre-existing knowledge and understanding of these speakers. Since the ability to use disfluencies to predict a speaker's intention to label a novel object emerges around children's second birthday (Kidd et al., 2012), children need to acquire certain knowledge and skills before they can make the most of these regular hesitations in the speech stream, if not for word learning, at least for real-time comprehension.

The work in this dissertation made important contributions to the literature on the real-time processing of speech disfluencies in children. Yet, they represent only a first step toward the goal of building a more comprehensive picture of how children process natural, everyday speech, which includes the comprehension of disfluencies. More work is needed to fully understand the role that certain features of speech disfluencies (e.g., lengthening of the speech stream, distinct determiners, the filled pause itself), in conjunction with characteristics of the speaker and listeners' perception of the speaker, influence the comprehension and word learning aspects of children's language development.

Lastly, the results from this dissertation can inform parents and educators about the role of spontaneous speech – including disfluencies – in language acquisition. Particularly for bilingual or non-native caregivers speaking to children, whether the filled pauses that they produce are in their native language or not does not seem to affect children's comprehension as they listen to speech, and their ability to use disfluencies to predict a speaker's intent to label a novel object. Studies of language acquisition are just beginning to study the disfluent nature of speech, despite how frequent speech disfluencies are in everyday conversation. The results presented here paint a complex picture – speech disfluencies may sometimes help children to identify a speaker's intended referent, but do not necessarily help them when faced with the complex task of forming a link between a novel word and this referent. More studies that consider the imperfect nature of speech are crucial. In addition to offering a more realistic portrayal of the linguistic challenges faced by young children, disfluent speech can inform us about the strategies that young listeners adopt to overcome these challenges as they embark on the great adventure of language acquisition.

References

Acton, E. K. (2011). On gender differences in the distribution of um and uh. *University of Pennsylvania Working Papers in Linguistics*, *17*(2), 2.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, *32*(1), 25–36. doi: 10.1023/A:1021980931292

Arnold, J. E., Kam, C. L. H., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 914. doi: 10.1037/0278-7393.33.5.914

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, *76*(1), 28-55. doi: 10.1353/lan.2000.0045

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new disfluency and reference resolution. *Psychological Science*, *15*(9), 578–582. doi: 10.1111/j.0956-7976.2004.00723.x

Arnold, J. E., & Tanenhaus, M. K. (2011). Disfluency effects in comprehension: How new information can become accessible. In E. Gibson & N. Perlmutter (Eds.), *The Processing and Acquisition of Reference.* Cambridge, MA: MIT Press. Doi: 10.7551/mitpress/9780262015127.003.0008

Astington, J. W., & Edward, M. J. (2010). The development of theory of mind in early childhood. *Social Cognition in Infancy*, *5*, 16.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. doi: 10.1016/j.tics.2009.12.006

Barr, D. J., & Seyfeddinipur, M. (2010). The role of fillers in listener attributions for speaker disfluency. *Language and Cognitive Processes*, *25*(4), 441–455. doi: 10.1080/01690960903047122

Beattie, G. W., & Butterworth, B. L. (1979). Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech*, *22*(3), 201–211. doi: 10.1177/002383097902200301

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

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*(1), 103-119.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39-53. doi: 10.1016/j.cognition.2012.08.008

Boersma, P., & Weenink, D. (2012). Praat (Version 5.1).

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, *44*(2), 123-147. doi: 10.1177/00238309010440020101

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

Bosker, H. R., Quené, H., Sanders, T., & de Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language*, *75*, 104–116. doi: 10.1016/j.jml.2014.05.004

Boudreault, M. C., Cabirol, É. A., Trudeau, N., Poulin-Dubois, D., & Sutton, A. (2007). Les inventaires MacArthur du développement de la communication: validité et données normatives préliminaires. *Revue canadienne d'orthophonie et d'audiologie*, *31*(1), 27–37.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, *44*(2), 274-296. doi: 10.1006/jmla.2000.2753

Brennan, S. E., & Williams, M. (1995). The feeling of another′ s knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*(3), 383-398. doi: 10.1006/jmla.1995.1017

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, *61*(1-2), 93-125. doi: 10.1016/s0010-0277(96)00719-6

Broen, P. A. (1972). The verbal environment of the language-learning child. ASHA Monographs, No. 17.

Brooker, I., & Poulin-Dubois, D. (2013). Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy*, *18*(s1), E46–E68. doi: 10.1111/infa.12027

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Developmental Psychology*, *38*(6), 958. doi: 10.1037//0012-1649.38.6.958

Brosseau-Liard, P. E., & Poulin-Dubois, D. (2014). Sensitivity to confidence cues increases during the second year of life. *Infancy*, *19*(5), 461–475. doi: 10.1111/infa.12056

Brosseau-Liard, P., Penney, D., Poulin-Dubois, D. (2015). Theory of mind selectively predicts preschoolers' knowledge-based selective learning. *British Journal of Developmental Psychology, 33*(4), 464-475. doi: 10.1111/bjdp.12107

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

Byers-Heinlein, K., Chen, K. H., & Xu, F. (2014). Surmounting the Tower of Babel: Monolingual and bilingual 2-year-olds' understanding of the nature of foreign language words. *Journal of Experimental Child Psychology*, *119*, 87–100. doi: 10.1016/j.jecp.2013.09.011

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. doi: 10.1002/dev.21167

Byers-Heinlein, K., Fennell, C. T., & Werker, J. F. (2013). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition*, *16*(1), 198-205. doi: 10.1017/S1366728912000417

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

Byers-Heinlein, K., Schott, E., Gonzalez-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.

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. doi: 10.1111/j.1467-7687.2009.00902.x

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, *128*(01), 407–416. doi: 10.1016/j.cognition.2013.05.010

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430. doi: 10.1037//0278-7393.27.6.1430

CIHR, NSERC, & SSHRC (Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada) (December 2014). *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. Accessed on May 16, 2017 from http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/

Clark, H. H., & Fox Tree, E. F. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111. doi: 10.1016/S0010-0277(02)00017-3

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, *37*(3), 201–242. doi: 10.1006/cogp.1998.0693

Clark, H. H., & Wilkes-Gibbs, D. (1990). Referring as a collaborative process. *Cognition*, *22*(1), 1-39. doi: 10.1016/0010-0277(86)90010-7

Colonnesi, C., Stams, G. J. J., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352-366. doi: 10.1016/j.dr.2010.10.001

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356-371. doi: 10.1016/j.cognition.2009.10.009

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584-1595. doi: 10.0.4.87/j.1467-8624.1990.tb02885.x

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)*,* 1637-1649. doi:10.1044/1092-4388

Corley, M., & Hartsuiker, R. J. (2011). Why um helps auditory word recognition: The temporal delay hypothesis. *PLoS One*, *6*(5), e19792. doi: 10.1371/journal.pone.0019792

Cutler, A. (1981). The reliability of speech error data. *Linguistics*, *19*(8), 561-82.

Davidson, D., Jergovic, D., Imami, Z., & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language*, *24*(1), 3-24. doi: 10.1017/s0305000996002917

Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology*, *92*(1), 25-45. doi: 10.1016/j.jecp.2005.03.007

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176. doi: 10.1126/science.7375928

Dink, J.W., & Ferguson, B. (2015). eyetrackingR: An R library for eye-tracking data analysis. Available at: http://www.eyetrackingr.com.

Duchateau, J., Laureys, T., Demuynck, K., & Wambacq, P. (2003). Handling disfluencies in spontaneous language models. *Language and Computers*, *47*, 39-50.

Duez, D. (1982). Silent and non-silent pauses in three speech styles. *Language and Speech*, *25*(1), 11–28. doi: 10.1177/002383098202500102

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*(6), 581-604. doi: 10.1016/s0149-7634(00)00025-7

Engelhardt, P. E., Nigg, J. T., & Ferreira, F. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function?. *Acta Psychologica*, *144*(2), 424-432. doi: 10.1016/j.actpsy.2013.08.002

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur-Bates Communicative Development Inventories*.

Fernald, A., Frank, M., Goodman, N., & Tenenbaum, J. (2009, January). Continuity of discourse provides information for word learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 31, No. 31).

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203-222. doi: 10.1111/j.1467-8624.2011.01692.x

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98. doi: 10.1037/0012-1649.42.1.98

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228-231. doi: 10.1111/1467-9280.00044

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104. doi: 10.1037//0012-1649.20.1.104

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

Finlayson, I. R., & Corley, M. (2012). Disfluency in dialogue: An intentional signal from the speaker?. *Psychonomic Bulletin & Review*, *19*(5), 921-928. doi: 10.3758/s13423-012-0279-x

Fischer, K. (2000). *From cognitive semantics to lexical pragmatics: The functional polysemy of discourse particles*. Walter de Gruyter.

Fox Tree, J. E. F. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *34*(6), 709–738. doi: 10.1006/jmla.1995.1032

Fox Tree, J. E. F. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, *29*(2), 320–326. doi: 10.3758/BF03194926

Fox Tree, J. E. F., & Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition*, *62*(2), 151-167. doi: 10.1016/s0010-0277(96)00781-0

Frank, I., & Poulin-Dubois, D. (2002). Young monolingual and bilingual children's responses to violation of the Mutual Exclusivity Principle. *International Journal of Bilingualism*, *6*(2), 125–146. doi: 10.1177/13670069020060020201

Friedrich, M., & Friederici, A. D. (2005). Semantic sentence processing reflected in the event-related potentials of one-and two-year-old children. *Neuroreport*, *16*(16), 1801–1804. doi: 10.1097/01.wnr.0000185013.98821.62

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513-543. doi: 10.1017.S0142716406060383

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition*, *6*(1), 1–15. doi: 10.1017/S1366728903001007

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178–186. doi: 10.1016/S1364-6613(00)01467-4

Grosjean, F. (1982). Life with two languages: An introduction to bilingualism. Harvard University Press.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23-B34. doi: 10.1016/s0010-0277(02)00186-5

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, *14*(3), 277-288. doi: 10.1177/002383097101400308

Heller, D., Arnold, J. E., Klein, N., & Tanenhaus, M. K. (2015). Inferring difficulty: Flexibility in the real-time processing of disfluency. *Language and Speech*, *58*(2), 190–203. doi: 10.1177/0023830914528107

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. doi: 10.1017/s0305000910000759

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). I. What Does it Take to Learn a Word? *Monographs of the society for research in child development*, *65*(3), 1-16.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., ... & Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, i-135.

Holowka, S., Brosseau-Lapré, F., & Petitto, L. A. (2002). Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language Learning*, *52*(2), 205-262. doi: 10.1111/0023-8333.00184

Hopper, T. (2016, July 18). 'Uh, certainly, I, uh': Does Justin Trudeau really say 'uh' all the time? Retrieved from http://nationalpost.com/news/canada/uh-certainly-i-uh-does-justin-trudeau-really-say-un-all-the-time

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. doi: 10.3758/s13428-015-0647-3

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. doi: 10.1080/15250000701795598

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*(2), 125–150. doi: 10.1111/j.1532-7078.2009.00009.x

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, *95*(1), 27-55. doi: 10.1016/j.jecp.2006.03.006

Hunter, M. A., Ross, H. S., & Ames, E. W. (1982). Preferences for familiar or novel toys: Effect of familiarization time in 1-year-olds. *Developmental Psychology*, *18*(4), 519. doi:

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

Jardak, A., Lew-Williams, C., & Byers-Heinlein, K. (in prep). *The impact of code switching on downstream word learning and recognition.*

Jescheniak, J. D., & Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824-843. doi: 10.1037/0278-7393.20.4.824

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548-567. doi: 10.1006/jmla.2000.2755

Johnson, E. K., & Zamuner, T. (2010). Using infant and toddler testing methods in language acquisition research. In Blom, E., Unsworth, S. (Eds.), Experimental methods in language acquisition research (pp. 73–94). Amsterdam, the Netherlands: John Benjamins.

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

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3-4), 159-207. doi: 10.1006/cogp.1999.0716

Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition*, *18*(4), 626–638. doi: 10.1017/S1366728914000364

Kidd, C., White, K. S., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science*, *14*(4), 925–934. doi: 10.1111/j.1467-7687.2011.01049.x

Kim, S., Kalish, C. W., & Harris, P. L. (2012). Speaker reliability guides children's inductive inferences about novel properties. *Cognitive Development*, *27*(2), 114–125. doi: 10.1016/j.cogdev.2011.10.004

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, *12*(1), 48–54. doi: 10.1111/j.1467-7687.2008.00742.x

Kucker, S. C. (2014). *The role of vocabulary knowledge and novelty biases in word learning: Exploring referent selection and retention in 18 to 24 month old children and associative models.* ProQuest Information & Learning, US.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing down fast mapping: Redefining the dynamics of word learning. *Child Development Perspectives*, *9*(2), 74-78. doi: 10.1111/cdep.12110

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2). doi: 10.1111/j.1467-7687.2006.00468.x

Labov, W. 1990. The intersection of sex and social class in the course of linguistic change. Language Variation and Change 2(2): 205-254. doi: 10.1017/s0954394500000338

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size and speed of word recognition in very young French–English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition*, 1-13. doi: 10.1017/S1366728916000833

Le Grezause, E. S. (2017). *Um and Uh, and the Expression of Stance in Conversational Speech* (Doctoral dissertation).

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*(3), 193–198. doi: 10.1111/j.1467-9280.2007.01871.x

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

Lundgren, K., Helm-Estabrooks, N., & Klein, R. (2010). Stuttering following acquired brain damage: A review of the literature. *Journal of Neurolinguistics*, *23*(5), 447-454. doi: 10.1016/j.jneuroling.2009.08.008

MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-Month-Olds' Phonotactic Knowledge Guides Their Word–Object Mappings. *Child Development*, *83*(4), 1129-1136. doi: 10.1111/j.1467-8624.2012.01764.x

MacKenzie, H., Graham, S. A., & Curtin, S. (2011). Twelve-month-olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science*, *14*(2), 249-255. doi: 10.1111/j.1467-7687.2010.00975.x

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53-59. doi: 10.1111/j.1467-7687.2007.00645.x

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3). doi: 10.1111/j.1467-7687.2007.00671.x

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language*, *37*(4), 817-840. doi: 10.1017/s0305000909990055

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121-157. doi: 10.1016/0010-0285(88)90017-5

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human perception and performance*, *15*(3), 576. doi: 10.1037//0096-1523.15.3.576

Mather, E., & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*, *14*(1), 60-76. doi: 10.1080/15250000802569702

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831. doi: 10.1037/a0029872

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, *7*(5), 491-497. doi: 10.1068/p070491

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227. doi: 10.1037/h0031564

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development, 16*(4), 495-500. doi: 10.1016/0163-6383(93)80007-u

Morin-Lessard, E., & Byers-Heinlein, K. (*under review*). *Uh* and *euh* signal novelty for monolinguals and bilinguals: Evidence from children and adults.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*(2), 136-142. doi: 10.1111/1467-7687.00263

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1-19. doi: 10.1006/jmla.2000.2698

Nilsson Björkenstam, K., Wirén, M., & Eklund, R. (2013). Disfluency in child-directed speech. In *XXVIth Annual Phonetics Meeting: Fonetik 2013, Linköping University, 12-13 June 2013, Linköping, Sweden* (pp. 57-60). Linköping University Electronic Press.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. doi: 10.1126/science.1107621

Orena, A. J., & White, K. S. (2015). I forget what that's called! Children's online processing of disfluencies depends on speaker knowledge. *Child Development*, *86*(6), 1701–1709. doi: 10.1111/cdev.12421

Owens, S. J., & Graham, S. A. (2016). Thee, uhh disfluency effect in preschoolers: A cue to discourse status. *British Journal of Developmental Psychology*, *34*(3), 388–401.doi: 10.1111/bjdp.12137

Owens, S. J., Thacker, J. M., & Graham, S. A. (2017). Disfluencies signal reference to novel objects for adults but not children. *Journal of Child Language*, 1-29. doi: 10.1017/S030500917000368

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278. doi: 10.1016/j.cortex.2015.04.014

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

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning*, *44*(4), 617-653. doi: 0.1111/j.1467-1770.1994.tb00633.x

Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tetreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, *28*(2), 453-496. doi: 10.1017/s0305000901004718

Petitto, L. A., Kovelman, I. (2003). The Bilingual Paradox: How signing-speaking bilingual children help us to resolve it and teach us about the brain’s mechanisms underlying all language acquisition. *Learning Languages*, 8:5–18.

Pomper, R., & Saffran, J. R. (2018). Familiar Object Salience Affects Novel Word Learning. *Child Development*. doi: 10.1111/cdev.13053

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en Español: Toward a typology of code-switching. *Linguistics, 18(7-8)*, 581–618. doi: 10.1515/ling.1980.18.7-8.581

Potter, C. E., Fourakis, E., Morin-Lessard, E., Byers-Heinlein, K., & Lew-Williams, C. (*under review).* Bilingual toddlers process mixed sentences differently in their two languages.

Poulin-Dubois, D., & Brosseau-Liard, P. (2016). The developmental origins of selective social learning. *Current Directions in Psychological Science*, *25*(1), 60-64. doi:

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development*, *77*(2), 266-280. doi: 10.1111/j.1467-8624.2006.00869.x

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*, *31*(1), 94–105. doi: 10.1080/23273798.2015.1077979

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*(1), 96-121. doi: 10.1016/j.cogpsych.2009.02.002

R Core Team (2017) R: A language and environment for statistical computing. Available at: https://www.R-project.org/

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word?. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2). doi: 10.1002/wcs.1421

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*(1), 94-104. doi: 10.1111/j.1467-8624.1998.tb06136.x

Segalowitz, N. (2009). *The Language Background Questionnaire*. Unpublished instrument, Department of Psychology, Concordia University, Montreal, QC, Canada.

Shriberg, E. (1996). Disfluencies in switchboard. In *Proceedings of International Conference on Spoken Language Processing,* Vol. 96, pp. 11-14. Philadelphia, PA, 3-6 October.

Shriberg, E. (2001). To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153-169. doi: 10.1017/s0025100301001128

Siri Team, (2017, October). *Hey Siri: An On-device DNN-powered Voice Trigger for Apple's Personal Assistant*. Retrieved from https://machinelearning.apple.com/2017/10/01/hey-siri.html

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*(1), 25–38. doi: 10.1006/jmla.1993.1002

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*(3), 238–299. doi: 10.1016/j.cogpsych.2004.03.001

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

Soderstrom, M., & Morgan, J. L. (2007). Twenty-two-month-olds discriminate fluent from disfluent adult-directed speech. *Developmental Science*, *10*(5), 641-653. doi: 10.1111/j.1467-7687.2007.00605.x

Stock, H. R., Graham, S. A., & Chambers, C. G. (2009). Generic language and speaker confidence guide preschoolers' inferences about novel animate kinds. *Developmental Psychology*, *45*(3), 884. doi: 10.1037/a0015359

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics*, *30*(4), 485–496. doi: 10.1016/s0378-2166(98)00014-9

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

Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018). When it is apt to adapt: Flexible reasoning guides children's use of talker identity and disfluency cues. *Journal of Experimental Child Psychology*, *167*, 314-327. doi: 10.1016/j.jecp.2017.11.008

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*(1/2), 61-82. doi: 10.1515/cogl.2001.012

Trudeau, N., Frank, I., & Poulin-Dubois, D. (1997). Inventaire MacArthur du Développement de la Communication: Mots et gestes. *Unpublished manuscript*.

Vignovic, J.A., & Thompson, L. F. (2010). Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation. *Journal of Applied Psychology, 95,* 265-276. doi: 10.1037/a0018628.

Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The Quarterly Journal of Experimental Psychology*, *48*(4), 822-848. doi: 10.1080/14640749508401419

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, *13*(6), 258-263. doi: 10.1016/j.tics.2009.03.006

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

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34 (6),* 1289–1309. doi: 10.1037/0012-1649.34.6.1289

Werker, J. F., & Fennell, C. T. (2004). Listening to sounds versus listening to words: Early steps in word learning. *Weaving a Lexicon*, 79-109.

Werker, J. F., Polka, L., & Pegg, J. E. (1997). The conditioned head turn procedure as a method for testing infant speech perception. *Early Development and Parenting*, *6*(34), 171–178. doi: 10.1002/(sici)1099-0917(199709/12)6:3/4<171::aid-edp156>3.3.co;2-8

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*(1), 49-63. doi: 10.1016/s0163-6383(84)80022-3

Werker, J. F., Yeung, H., H., & Yoshida, K. A. (2012). How do infants become experts in native-speech perception? *Current Directions in Psychological Science, 21*(4), 221-226. doi: 10.1177/0963721412449459

White, K. S., Nilsen, E. S, & Riemersma, T. (*under review*). That's thee, uuh blicket! How does disfluency affect children's word learning? *First Language*.

Wieling, M., Grieve, J., Bouma, G., Fruehwald, J., Coleman, J., & Liberman, M. (2016). Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change*, *6*(2), 199-234. doi: 10.1163/22105832-00602001

Yairi, E., Gintautas, J., & Avent, J. R. (1981). Disfluent speech associated with brain damage. *Brain and Language*, *14*(1), 49-56. doi: 10.1016/0093-934x(81)90063-8

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory of mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development*, *17*(5), 683-698. doi:10.1080/15248372.2015.1086771

Yow, W. Q., & Markman, E. M. (2011). Bilingualism and children's use of paralinguistic cues to interpret emotion in speech. *Bilingualism: Language and Cognition*, *14*(4), 562–569. doi: 10.1017/S1366728910000404

Yow, W. Q., & Markman, E. M. (2016). Children increase their sensitivity to a speaker's nonlinguistic cues following a communicative breakdown. *Child Development*, *87*(2), 385-394. doi: 10.1111/cdev.12479

Appendix A: Consent Forms

CONSENT TO PARTICIPATE IN THE MONOLINGUAL AND BILINGUAL DEVELOPMENT PROJECT

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

A. PURPOSE

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

B. PROCEDURES

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

C. RISKS AND BENEFITS

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

D. CONDITIONS OF PARTICIPATION

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

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

CHILD'S NAME (please print) ____

PARENT'S NAME (please print)

____

SIGNATURE ____

DATE ____

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

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

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

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

Baby ID: ____________ Researcher: _________________________

CONSENT TO PARTICIPATE IN THE MONOLINGUAL AND BILINGUAL DEVELOPMENT PROJECT

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

A. PURPOSE

I understand that the purpose of the research is to investigate how individuals develop their language and conceptual skills.

B. PROCEDURES

I understand that the present study is approximately 30 minutes long. During this time, I will see an audio-visual presentation including one or more of the following: written words, language sounds, non-language sounds, colourful pictures, or a live interaction with a researcher. I may also be asked to respond verbally or via a button press, or to simply watch a presentation. My reactions throughout the study may be recorded on video and/or via an eye tracker, and will be kept by the researcher for future reference. I will also be asked to complete several questionnaires about my language knowledge and background. Data will be stored in a secure location at Concordia University, and participants will only be identified by code number to protect confidentiality.

C. RISKS AND BENEFITS

I understand that there are no known risks to participation in this study. As a thank you for my participation, I will be offered $5 or credit in the Psychology Department Participant Pool.

D. CONDITIONS OF PARTICIPATION

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

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

NAME (please print) ____

SIGNATURE ____

DATE ____

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

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

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

If at any time you have questions about your rights as a research participant, please contact the Research Ethics and Compliance Advisor, Concordia University, at (514)-848-2424 x7481 or by email at ethics@alcor.concordia.ca

Participant # ____________ Researcher: _________________________

Appendix B : Language Exposure Questionnaire

Language Exposure Questionnaire for Infants

A structured interview (do not hand directly to parents)

Baby_ID: ________________ Today's date (MM/DD/YYYY):____________________________

Study_ID: ________________

Parent completing questionnaire: ________________________

Study Name: ______________

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

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

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

DAYCARE: Does [child's name] attend regular childcare, such as daycare? **YES NO**
If yes, since what age: _________
What language(s) are spoken? [Note details] ____

TRIPS: Has [child's name] ever been on a trip of 1 month or more where his/her language exposure would have changed? **YES NO**

Day-in-the-life estimate

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

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

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

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

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

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

__________ __________ __________ __________

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

Overall estimate (Average of <u>Global estimate</u> and day-in-the-life <u>% Cumulative exposure</u>).
* **Monolingual exception:** If % cumulative exposure >90%, use % cumulative without averaging.

Language:				
Percent:				
Age Acquisition:				

Appendix C : Demographic Questionnaire

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

Concordia Infant Research Laboratory Participant Information

Child's Name: ____
First Last
Child's Date of Birth: ___________________ Child's Gender: ☐ M ☐ F
MM / DD / YY

Basic Family Information
Parent A's Full Name: ____ ☐ M ☐ F
First Last
Parent B's Full Name: ____ ☐ M ☐ F
First Last
Address (including postal code):
____
____

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

E-mail: ______________________________________

Does your child have any siblings?

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

Languages Spoken in the Home and at Childcare

What percent of the time does your baby hear English? _________ %
What percent of the time does your baby hear French? _________ %
What percent of the time does your baby hear another language? _________ %
Please specify this language:________________________

Has the child lived/vacationed in any country where s/he would hear a language other than English or French? ☐ Yes ☐ No
If yes, please detail (when, where, and for how long?) _________________________________

Health History

What was your child's birth weight? __ __ lbs __ __ oz OR __ __ __ __ grams
How many weeks was your pregnancy? ____________weeks

Were there any complications during the pregnancy? ☐ Yes ☐ No
If yes please detail ____

Has your child had any major medical problems?
If yes please detail____

Does your child have any hearing or vision problems?
If yes please detail____

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

Has your child had any ear infections in the past? ☐ Yes ☐ No
If yes at which ages____

Does your child have a cold today? ☐ Yes ☐ No
If yes, does he/she have pressure/pain in ears (if known)? ☐ Yes ☐ No
Is there any other relevant information we should know (health or language-related)?

Has another university contacted you to participate in one of their studies? ☐ Yes ☐ No
If yes, which university? _________________________________

Family and Child Background Information (optional)

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

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

____
____
______

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

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

____
____
______

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

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

____
____
______

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

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

____
____
______

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

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

____

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

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

____

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

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

____

Appendix D : Language Background Questionnaire

UNIVERSITÉ Concordia UNIVERSITY

LANGUAGE BACKGROUND QUESTIONNAIRE

Name: ________________________________ Date: ________________________

Age: ___________________ Sex: M ___ F___

1. If you are a student:

   What is your field of study? ____

   What degree are you pursuing? College (Cégep/Diploma/etc.)___ Bachelor___

MA/PhD/etc.___

2. Where were you born? City:___________________ Country: _______________________

3. What do you consider to be your **first learned language**?

   English ___ French ___ Other (specify) ________________

4. What do you consider to be your **second learned language**?

   English ___ French ___ Other (specify) ________________

5. At what age did you learn your **second language**? Age of ________ years old

6. What language do you consider your dominant language (the language you are most comfortable in)?

   English ___ French ___ Other (specify) ________________

7. What language do you speak at home now? _______________________________________

8. What is the first language of your: Mother? ____________ Father? ________________

9. What was the language of instruction of the school you attended? (Check all appropriate):

- Preschool:	English___	French___	French Immersion___	Other (specify)______
- Elementary school:	English___	French___	French Immersion___	Other (specify)______
- Middle/High school:	English___	French___	French Immersion___	Other (specify)______
- College/Cégep/Diploma:	English___	French___		Other (specify)______
- University:	English___	French___		Other (specify)______

10. If you are not currently a student, what is the highest level of education you have completed:

    High school___ College/Cégep/Diploma___

    University: Bachelor___ University: MA/PhD/etc.___

11. Did you receive second/foreign language instruction at any of the levels listed below
YES ___ NO ____

If YES, specify each language and for how long, starting with your main second language.

MAIN SECOND/FOREIGN LANGUAGE: ___________________________

- Elementary School: less than 1 year ___ 1-2 years ___ more than 2 years ___
- Middle/High School: less than 1 year ___ 1-2 years ___ more than 2 years ___
- College/Cégep/Diploma/University: less than 1 year ___ 1-2 years ___ more than 2 years ___
- Other: less than 1 year ___ 1-2 years ___ more than 2 years ___

Please specify: ____

THIS LANGUAGE (if any): ____

Any other special learning experiences (e.g., intensive French in Grade 6; long visit to France):

____

12. Do you have any visual impairment NOT corrected by wearing glasses or contact lenses? Yes ___ No ___

13. Do you have a known hearing impairment? Yes ___ No ___

14. Do you have a known reading or attention disability? Yes ___ No ___

15. What percentage of your interactions are in (total = 100%):

English __ %? French __%? Other ___%?

16. Please rate your level of ability for each of the four skills listed below by using the following rating scheme and circling the appropriate number in the boxes below:

1 = no ability at all **2** = elementary **3** = moderate **4** = very good **5** = fluent ability

Language	Speaking	Reading	Writing	Listening
English	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
French	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Other ____________	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

SPEAKING FRENCH, YOUR SECOND LANGUAGE

Please think about how well you can **SPEAK** French.

Now complete STEP 1.

	STEP 1
	First, read all the descriptions shown in this table and try to find **the one description** that best describes your **ability to speak French**. Next, place a check mark "4" in the box "**o**" next to that one description that best indicates your ability to **<u>speak</u>** French.
o	I can express myself in all or almost all contexts, using all or nearly all expressions that native speakers typically use.
o	I can express myself on a large number of unfamiliar topics, although I don't always know the expressions that native speakers would typically use.
o	I can only speak about a small number of familiar topics.
o	I am limited to saying only simple things, such as asking for directions or answering short questions.
o	I cannot express very much at all in the language.

Now complete STEP 2.

STEP 2
Read the "A" and "B" boxes below. Then circle "A" or "B" to indicate how easy it is for you to **<u>speak</u>** French at the level you selected above.

A	At the level I selected, I generally speak more or less fluently and at a normal rate.	**B**	At the level I selected, I sometimes hesitate, have to search noticeably for words, or try to avoid making errors, or speak more slowly than I usually do.

UNDERSTANDING SPOKEN FRENCH, YOUR SECOND LANGUAGE

Please think about how well you can **UNDERSTAND** spoken French.

Now complete STEP 1.

	STEP 1
	First, read all the descriptions shown in this table and try to find **the one description** that best describes your **ability to understand spoken French**. Next, place a check mark "4" in the box "**o**" next to that one description that best indicates your ability to **understand** spoken French.
o	I can understand native speakers in all or almost all contexts, including nearly all the expressions that native speakers typically use.
o	I can understand native speakers on a large number of unfamiliar topics, even though I may not always understand every expression they use.
o	I am limited to understanding native speakers only when they talk about a small number of familiar topics.
o	I can understand native speakers only when they talk about simple things, such as when they give directions or ask short questions.
o	I cannot understand others very much at all in the language.

Now complete STEP 2.

STEP 2
Read the "A" and "B" boxes below. Then circle "A" or "B" to indicate how easy it is for you to **understand** spoken French at the level you selected above.

A	At the level I selected, I generally understand speakers who speak fluently and at a normal speed.	**B**	At the level I selected, I sometimes have to ask people to slow down and repeat, or to speak more clearly, or to explain the meanings of some words.

UNDERSTANDING SPOKEN ____________________, YOUR THIRD LANGUAGE

Please think about how well you can **UNDERSTAND** spoken ____________________.

Now complete STEP 1.

	STEP 1
	First, read all the descriptions shown in this table and try to find **the one description** that best describes your **ability to understand spoken _____________**. Next, place a check mark "4" in the box "**o**" next to that one description that best indicates your ability to **understand** spoken _____________.
o	I can understand native speakers in all or almost all contexts, including nearly all the expressions that native speakers typically use.
o	I can understand native speakers on a large number of unfamiliar topics, even though I may not always understand every expression they use.
o	I am limited to understanding native speakers only when they talk about a small number of familiar topics.
o	I can understand native speakers only when they talk about simple things, such as when they give directions or ask short questions.
o	I cannot understand others very much at all in the language.

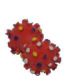
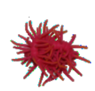
Now complete STEP 2.

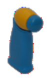
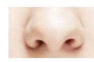
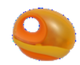
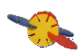
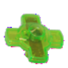
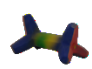
STEP 2
Read the "A" and "B" boxes below. Then circle "A" or "B" to indicate how easy it is for you to **understand** spoken _____________ at the level you selected above.

A	At the level I selected, I generally understand speakers who speak fluently and at a normal speed.	**B**	At the level I selected, I sometimes have to ask people to slow down and repeat, or to speak more clearly, or to explain the meanings of some words.

Appendix E : Pictures of Familiar-Novel Object Pairs in Manuscript 1

Familiar-novel object pairs with their respective novel and familiar objects in English and in French

English Order		French Order		
Familiar	Novel	Det.	Familiar	Novel
				
book	semp	le	bas [sock]	sèpe
				
bread	mog	la	bouche [mouth]	taque
				
cheese	gorp	la	chaise [chair]	moube
				
cookie	posha	le	chapeau [hat]	brussin
				
cow	sarl	la	cuillère [spoon]	dimotte
				
cup	blat	la	dent [tooth]	shème
				
dog	koob	le	lait [milk]	fide
				
doll	rel	le	lit [bed]	juve

				
frog	dak	la	main [hand]	froise
				
key	mip	la	maison [house]	télue
				
pig	juff	le	nez [nose]	migue
				
plate	zel	le	pied [foot]	vuche
				
stroller	voopi	la	pomme [apple]	kète
				
toothbrush	glindle	la	porte [door]	tine
				
truck	shob	le	soulier [shoe]	pafli
				
window	teeba	le	verre [glass]	triffe

Note. Images of familiar objects were selected from Google, and images of novel objects were selected from the NOUN-2 Database (Horst & Hout, 2016).

Appendix F: Discrimination Task Sheet

Dis-Adults – Discrimination Task

In this task you will hear either "uh" (English hesitation) or "euh" (French hesitation) on each trial. Please circle which one you hear for each trial (circle only ONE answer).

	English	French
1.	uh	euh
2.	uh	euh
3.	uh	euh
4.	uh	euh
5.	uh	euh
6.	uh	euh
7.	uh	euh
8.	uh	euh
9.	uh	euh
10.	uh	euh
11.	uh	euh
12.	uh	euh
13.	uh	euh
14.	uh	euh
15.	uh	euh
16.	uh	euh

	English	French
17.	uh	euh
18.	uh	euh
19.	uh	euh
20.	uh	euh
21.	uh	euh
22.	uh	euh
23.	uh	euh
24.	uh	euh
25.	uh	euh
26.	uh	euh
27.	uh	euh
28.	uh	euh
29.	uh	euh
30.	uh	euh
31.	uh	euh
32.	uh	euh